RESEARCH   **Open Access**

Joint CFO and DOA estimation for multiuser OFDMA uplink

Weile Zhang[1*], Qinye Yin[1] and Feifei Gao[2,3]

Abstract

In this article, we develop a new subspace-based multiuser joint carrier frequency offset (CFO) and direction-of-arrival (DOA) estimation scheme for orthogonal frequency division multiple access uplink transmissions. We leverage multi-antenna at the receiver and consider that the signals transmitted by each user arrive at the receiving antenna array from multiple DOAs after bouncing from both surrounding and far scatterers. The rank reduction approach is then exploited to estimate the multiple CFOs and DOAs. Specifically, for each user, after the CFO estimation from one-dimensional search, its multiple DOAs can be obtained simultaneously via polynomial rooting. The proposed method supports generalized subcarrier assignment scheme and fully loaded transmissions. Both performance analysis and numerical results are provided to corroborate the proposed studies.

Keywords: carrier frequency offset (CFO), direction-of-arrival (DOA), orthogonal frequency division multiple access (OFDMA)

Introduction

As has widely been studied in recent years [1-3], orthogonal frequency division multiple access (OFDMA) is deemed as a promising technique for next-generation multiuser wireless communications. The performance of OFDMA, however, is sensitive to multiple carrier frequency offsets (CFOs) introduced by the mismatch of the transceiver oscillators or the Doppler effect. In multiuser scenarios, the non-zero CFOs lead to both inter-carrier interference and multiple-access interference, which could severely degrade the system performance.

The CFO estimation scheme for OFDMA uplink transmissions has intensively been investigated in the past few years. Using the frequency domain embedded pilot symbols, an iterative CFO estimation approach was described in [4] for tile structure-based OFDMA transmission [5]. The CFOs can also be estimated from the maximum likelihood (ML) approach by transmitting training sequences from each user, but with very high complexity. The alternating-projection algorithm was introduced in [6] to replace the multi-dimensional search with a sequence of one-dimensional (1D) searches. An improved approach

was later proposed in [7,8] to further reduce the complexity of [6] by using the divide-and-update frequency estimator. An interesting alternative to avoid the ML multi-dimensional search is to use the mean likelihood estimator combined with the importance sampling technique [9,10]. Another complexity-reduced CFO estimator was reported in [11] by approximating the inverse of a CFO-dependent matrix with that of a predetermined matrix.

Blind CFO estimation methods, on the other hand, were also developed to improve the bandwidth efficiency. The CFOs can be computed by looking for the position of null subcarriers within the signal bandwidth in the system for subband subcarrier assignment scheme (SAS) [12]. A frequency estimation scheme for uplink OFDMA with interleaved SAS that exploits the periodic structure of the signals from each user has been reported in [13], where the subspace estimation theory was utilized, which makes the scheme similar to the multiple signal classification technique [14]. Based on the observation of [13], several advancements have been proposed later [15,16]. Despite their good performance, both [13] and its variations [15,16] are only applicable for interleaved SAS and cannot be used for generalized SAS. Moreover, they must reserve null subcarriers or a much longer cyclic prefix

* Correspondence: wlzhang1984@gmail.com
[1]MOE Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, P. R. China
Full list of author information is available at the end of the article

(CP) to construct the noise space, which reduces the bandwidth efficiency.

More recently, several CFO estimation schemes have been developed for the OFDMA systems by leveraging multi-antenna at the receiver. For instances, a CFO estimation scheme for interleaved OFDMA/space division multiple access uplink systems was developed in [17] to support spatially separated users and to maximize the channel throughput. Another several schemes were proposed in [18,19] to support generalized SAS as well as fully loaded transmissions. They adopted the estimation of signal parameters via rotational invariance technique (ESPRIT)-like approach and exploited the direction-of-arrival (DOA) information to separate the signals from different users.

In this article, we develop a new subspace-based multiuser joint CFO and DOA estimation scheme for OFDMA uplink transmissions. We leverage multi-antenna at the receiver and consider that the signals transmitted by each user arrive at the receiver's antenna array from multiple DOAs, after bouncing from both surrounding and far scatterers [20]. The multiple CFOs and DOAs are then derived by a rank-reduction approach. Specifically, for each user, after the CFO estimation using 1D search, its multiple DOAs can also be obtained simultaneously by polynomial rooting, which is one unique property of our scheme. In summary, the main contributions of this article include the following:

1. With the consideration of multi-cluster channels, we design a new joint CFO and DOA estimation method for multiuser OFDMA uplink. The proposed method supports generalized SAS and fully loaded transmissions.
2. We provide the theoretical performance analysis of our method in terms of both CFO and DOA estimation.
3. Compared with [18,19], the simulation results demonstrate that our method not only has the advantage of being applicable to multi-cluster channels, but also can obtain much better performance in single cluster channels.

Notations: Superscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $[\cdot]^\dagger$, and $E[\cdot]$ represent conjugate, transpose, Hermitian, pseudo inverse, and expectation, respectively; $\mathbf{j} = \sqrt{-1}$ is the imaginary unit; $||\mathbf{X}||$ denotes the Frobenius norm of $\mathbf{X}$, and diag($\cdot$) is a diagonal matrix with main diagonal $(\cdot)$; The kronecker product is denoted by $\otimes$; The component-wise product is denoted by $\circ$; $\mathbf{I}_N$ denotes the $N \times N$ identity matrix and $\mathbf{1}_N$ denotes the $1 \times N$ matrix with all entries being 1; Matlab matrix representations are adopted, for example, $\mathbf{X}(r_1 : r_2, c_1 : c_2)$ denotes the submatrix of $\mathbf{X}$ with the rows from r_1 to r_2 and the columns from c_1 to c_2.

System model

We consider a multiuser OFDMA system with K users, N subcarriers. The base station (BS) is equipped with a uniform linear array (ULA) with M antennas, which is elevated above the rooftop. All subcarriers are sequentially indexed with $\{0, 1, \ldots, N - 1\}$. Assume that the channel between each user and the receiver is composed of N_{cl} clusters ($N_{cl} \geq 1$). The multipath components in each cluster exhibit similar DOAs. Among the total N_{cl} clusters, one cluster is called surrounding cluster that corresponds to the scatterers located around each user, and the remaining $N_{cl} - 1$ clusters, called far clusters, correspond to high-rise buildings in urban environments and hills/mountains in rural environments [20,21]. As the BS is deployed above its surrounding scatterers, following [18,19], we further approximate that the multipath components from one cluster have a single DOA. Here we should note that the works in [18,19] considered only the surrounding cluster, but ignored the existence of far clusters. However, as has been reported in [20], in the typical Urban environment, the fractions of the cases with two and three clusters are 9 and 4%, respectively. The fractions are even higher in the bad Urban environment, which are given by 28 and 45%, respectively. We should note that the methods developed in [18,19] may not be applicable to these multi-cluster scenarios.

Denote $\chi = d/\lambda$, where d is the antenna spacing of the ULA, and λ is the radio wavelength. Assume that in the gth block, the multipath channel components between the ith cluster of the kth user and the reference antenna (1st) of ULA can be modeled by a length-L_p vector

$$\mathbf{h}_{1,i,g}^{(k)} = \left[h_{i,g}^{(k)}(0), h_{i,g}^{(k)}(1), \ldots, h_{i,g}^{(k)}(L_p - 1) \right]^T. \quad (1)$$

We assume that the entries of $\mathbf{h}_{1,i,g}^{(k)}$ are independent Gaussian variables with variance $1/L_p$ such that the expectation of the channel vector norm is 1, i.e., $E[\| \mathbf{h}_{1,i,g}^{(k)} \|] = 1$.

Let $\varphi_i^{(k)}$ denote the DOA of the ith cluster of the kth user and then the multipath channel components between the ith cluster of the kth user and the mth antenna of ULA can be expressed as

$$\mathbf{h}_{m,i,g}^{(k)} = a_{m,i}^{(k)} \cdot \mathbf{h}_{1,i,g}^{(k)} \quad (2)$$

where $a_{m,i}^{(k)} = e^{\mathbf{j} 2\pi \chi (m-1) \cos \varphi_i^{(k)}}$. Correspondingly, its frequency domain channel is given by

$$
\begin{aligned}
\mathbf{H}_{m,i,g}^{(k)} &= \sqrt{N} \mathbf{F} \cdot [(\mathbf{h}_{m,i,g}^{(k)})^T, \mathbf{0}_{1 \times (N - L_p)}]^T \\
&= a_{m,i}^{(k)} \cdot [H_{i,g}^{(k)}(0), H_{i,g}^{(k)}(1), \ldots, H_{i,g}^{(k)}(N - 1)]^T,
\end{aligned} \quad (3)
$$

where $\mathbf{F}$ stands for the $N \times N$ DFT matrix with its (i, j) entry $\mathbf{F}(i, j) = \frac{1}{\sqrt{N}}e^{-j2\pi\frac{(i-1)(j-1)}{N}}$. We denote the normalized CFO of the kth user by $\xi^{(k)} = \Delta f^{(k)}/\Delta f$, where Δf is the subcarrier spacing and $\Delta f^{(k)}$ is the CFO of the kth user. We assume $\xi^{(k)} \in$ (-0.5, 0.5). Denote the number and the index set of the subcarriers allocated to the kth user by N_k and $C^{(k)}$, where

$$C^{(k)} = \{c_1^{(k)}, c_2^{(k)}, \ldots, c_{N_k}^{(k)}\}, \quad \sum_{k=1}^{K} N_k = N_{\text{sum}} \leq N. \quad (4)$$

Let $\mathbf{s}_g^{(k)} = [s_{1,g}^{(k)}, s_{2,g}^{(k)}, \ldots, s_{N_k,g}^{(k)}]^T$ be the modulated symbols of the kth user in the gth block. In the noise-free environment, the received time-domain signal components after removing CP from the kth user at the mth antenna can be expressed as

$$\gamma_{m,g}^{(k)}(n) = \frac{1}{\sqrt{N}}\sum_{p=1}^{N_k}e^{j\frac{2\pi}{N}(c_p^{(k)}+\xi^{(k)})n}\left(\sum_{i=1}^{N_{\text{cl}}}a_{m,i}^{(k)}H_{i,g}^{(k)}(c_p^{(k)})\right)s_{p,g}^{(k)},(5)$$

where the term in the bracket stands for the composition frequency-domain channel response at the $c_p^{(k)}$th subcarrier of the kth user resulting from total N_{cl} clusters. Then the overall received signal from K users at the mth antenna can be expressed as

$$\gamma_{m,g}(n) = \sum_{k=1}^{K}\gamma_{m,g}^{(k)}(n) = \frac{1}{\sqrt{N}}\sum_{k=1}^{K}\sum_{p=1}^{N_k}\sum_{i=1}^{N_{\text{cl}}}a_{m,i}^{(k)}X_{i,p,g}^{(k)}e^{j2\pi\Theta_p^{(k)}n}\,(6)$$

where $X_{i,p,g}^{(k)} = H_{i,g}^{(k)}(c_p^{(k)})s_{p,g}^{(k)}$ and $\Theta_p^{(k)} = \frac{c_p^{(k)}+\xi^{(k)}}{N}$ denote the effective CFO on the $c_p^{(k)}$th subcarrier of the kth user.

Stacking the received signals from all M antenna elements at the nth sample, we obtain the following space-domain snapshot vector

$$\gamma_{n,g} = \left[\gamma_{1,g}(n), \gamma_{2,g}(n), \ldots, \gamma_{M,g}(n)\right]^T. \quad (7)$$

Define the Vandermonde vector

$$\mathbf{a}_i^{(k)} = \left[a_{1,i}^{(k)}, a_{2,i}^{(k)}, \ldots, a_{M,i}^{(k)}\right]^T \quad (8)$$

which reflects the DOA of the ith cluster of the kth user, and obtain the corresponding Vandermonde matrix

$$\mathbf{a}^{(k)} = \left[\mathbf{a}_1^{(k)}, \mathbf{a}_2^{(k)}, \ldots, \mathbf{a}_{N_{\text{cl}}}^{(k)}\right] \quad (9)$$

by collecting N_{cl} Vandermonde vectors. Considering the noise item, we can then rewrite γ_{ng} in the following matrix form

$$\gamma_{n,g} = \frac{1}{\sqrt{N}}\mathbf{\Omega}\mathbf{\Phi}^n\mathbf{X}_g + \mathbf{n}_{n,g}, \quad (10)$$

where

$$\mathbf{\Omega} = \left[\mathbf{1}_{N_1}\otimes a^{(1)}, \mathbf{1}_{N_2}\otimes a^{(2)}, \ldots, \mathbf{1}_{N_K}\otimes a^{(K)}\right], \quad \mathbf{X}_{p,g}^{(k)} = \left[X_{1,p,g}^{(k)}, X_{2,p,g}^{(k)}, \ldots, X_{N_{\text{cl}},p,g}^{(k)}\right]^T,$$

$$\mathbf{X}_g^{(k)} = \left[(\mathbf{X}_{1,g}^{(k)})^T, (\mathbf{X}_{2,g}^{(k)})^T, \ldots, (\mathbf{X}_{N_k,g}^{(k)})^T\right]^T, \quad \mathbf{X}_g = \left[(\mathbf{X}_g^{(1)})^T, (\mathbf{X}_g^{(2)})^T, \ldots, (\mathbf{X}_g^{(K)})^T\right]^T,$$

$$\mathbf{\Phi}^{(k)} = \text{diag}(e^{j2\pi\Theta_1^{(k)}}, e^{j2\pi\Theta_2^{(k)}}, \ldots, e^{j2\pi\Theta_{N_k}^{(k)}})\otimes I_{N_{\text{cl}}}, \quad \mathbf{\Phi} = \text{diag}(\mathbf{\Phi}^{(1)}, \mathbf{\Phi}^{(2)}, \ldots, \mathbf{\Phi}^{(K)}),$$

and $\mathbf{n}_{n,g}$ is a length-M additive white Gaussian noise (AWGN) vector with variance matrix $\sigma_n^2\mathbf{I}_M$ at the nth sample in the gth block.

Joint CFO and DOA estimation
Properties of the subspace
Stacking L ($L \leq N$) continuous space-domain snapshot vectors from the nth to the $(n+L-1)$th sample time, we obtain

$$\gamma_g|_n^{n+L-1} = \left[(\gamma_{n,g})^T, (\gamma_{n+1,g})^T, \ldots, (\gamma_{n+L-1,g})^T\right]^T$$
$$= \mathbf{A}\mathbf{\Phi}^n\mathbf{X}_g + \mathbf{N}_g|_n^{N+L-1} \quad (11)$$

where

$$\mathbf{A} = \frac{1}{\sqrt{N}}\left[(\mathbf{\Omega})^T, (\mathbf{\Omega}\mathbf{\Phi})^T, \ldots, (\mathbf{\Omega}\mathbf{\Phi}^{L-1})^T\right]^T, \quad \mathbf{N}_g|_n^{n+L-1} = \left[\mathbf{n}_{n,g}^T, \mathbf{n}_{n+1,g}^T, \ldots, \mathbf{n}_{n+L-1,g}^T\right]^T.$$

The effect of the parameter L will be discussed later. Afterwards, by defining

$$\mathbf{b}_p^{(k)} = \frac{1}{\sqrt{N}}[1, e^{j2\pi\Theta_p^{(k)}}, \ldots, e^{j2\pi(L-1)\Theta_p^{(k)}}]^T, \quad \mathbf{B}^{(k)} = [\mathbf{b}_1^{(k)}, \mathbf{b}_2^{(k)}, \ldots, \mathbf{b}_{N_k}^{(k)}],$$

we can rewrite $\mathbf{A}$ as

$$\mathbf{A} = \left[\mathbf{B}^{(1)}\otimes\mathbf{a}^{(1)}, \mathbf{B}^{(2)}\otimes\mathbf{a}^{(2)}, \ldots, \mathbf{B}^{(K)}\otimes\mathbf{a}^{(K)}\right]. \quad (12)$$

We obtain the correlation matrix of $\gamma_g|_n^{n+L-1}$ as follows

$$\mathbf{R}_\gamma = E\left[\gamma_g|_n^{n+L-1}(\gamma_g|_n^{n+L-1})^H\right] = \mathbf{A}\mathbf{R}_{XX}\mathbf{A}^H + \sigma_n^2\mathbf{I}_{ML}, \quad (13)$$

where
$$\mathbf{R}_{XX} = \frac{1}{N-L+1}\sum_{n=0}^{N-L}\mathbf{\Phi}^nE[\mathbf{X}_g\mathbf{X}_g^H](\mathbf{\Phi}^n)^H = \sigma_s^2\mathbf{I}_{N_{\text{cl}}N_{\text{sum}}} \quad \text{with}$$
σ_s^2 being the average power of the transmitted signals. Then, we have $\mathbf{R}_\gamma = \sigma_s^2\mathbf{A}\mathbf{A}^H + \sigma_n^2\mathbf{I}_{ML}$. In practice, using successive L_s OFDMA blocks, this correlation matrix can be approximatd by

$$\widehat{\mathbf{R}}_\gamma = \frac{1}{(N-L+1)L_s}\sum_{g=1}^{L_s}\sum_{n=0}^{N-L}\gamma_g|_n^{n+L-1}(\gamma_g|_n^{n+L-1})^H.(14)$$

Afterwards, we assume the matrix $\mathbf{A}$ is tall and has full column rank, and the corresponding discussion will be presented later. Performing singular value decomposition (SVD) on $\mathbf{R}_\gamma$ gives.

$$\mathbf{R}_\gamma = [\mathbf{U}_\gamma, \ \mathbf{V}_\gamma] \Sigma_\gamma [\mathbf{U}_\gamma, \ \mathbf{V}_\gamma]^H, \tag{15}$$

where $\mathbf{U}_\gamma$ and $\mathbf{V}_\gamma$ represent the $(N_{cl}N_{sum})$-dimensional signal space and $(ML - N_{cl}N_{sum})$-dimensional noise space matrices, respectively. We define the following length-L parameterized Vandermonde vector with respect to ξ:

$$\mathbb{B}_p^{(k)}(\xi) = \frac{1}{\sqrt{N}} \left[1, \ e^{j2\pi \frac{c_p^{(k)}+\xi}{N}}, \ \dots, \ e^{j2\pi \frac{(L-1)(c_p^{(k)}+\xi)}{N}} \right]^T, \tag{16}$$

where $\xi \in$ (-0.5, 0.5). Clearly, there holds $\mathbf{b}_p^{(k)} = \mathbb{B}_p^{(k)}(\xi^{(k)})$. For notational convenience, we denote $\mathbf{0}$ as the all-zero matrix with appropriate dimension. Lemma 1 gives the key properties to design our joint estimator:

Lemma 1: When the matrix $\underline{\mathbf{A}}$ has full column rank, then for a non-zero length-M vector ω, there holds

$$\left(\mathbb{B}_p^{(k)}(\xi^{(k)}) \otimes \omega \right)^H \mathbf{V}_\gamma = \begin{cases} = \mathbf{0}, & \omega \in \mathrm{Span}(a^{(k)}), \\ \neq \mathbf{0}, & \omega \notin \mathrm{Span}(a^{(k)}), \end{cases} \tag{17}$$

and

$$\left(\mathbb{B}_p^{(k)}(\xi) \otimes \omega \right)^H \mathbf{V}_\gamma \neq \mathbf{0}, \quad \xi \neq \xi^{(k)}, \tag{18}$$

where $\underline{\mathbf{A}}$ is the first $M(L-1)$ rows of $\mathbf{A}$ which can be expressed as

$$\underline{\mathbf{A}} = \left[\underline{\mathbf{B}}^{(1)} \otimes a^{(1)}, \ \underline{\mathbf{B}}^{(2)} \otimes a^{(2)}, \ \dots, \ \underline{\mathbf{B}}^{(K)} \otimes a^{(K)} \right],$$
$$\underline{\mathbf{B}}^{(k)} = \left[\underline{\mathbf{b}}_1^{(k)}, \ \underline{\mathbf{b}}_2^{(k)}, \ \dots, \ \underline{\mathbf{b}}_{N_k}^{(k)} \right], \quad \underline{\mathbf{b}}_p^{(k)} = \frac{1}{\sqrt{N}} [1, \ e^{j2\pi \Theta_p^{(k)}}, \ \dots, \ e^{j2\pi(L-2)\Theta_p^{(k)}}]^T.$$

Proof. See Appendix 1. □

Parameters estimation
CFO estimation
For any non-zero length-M vector ω, we know

$$\sum_{p=1}^{N_k} (\mathbb{B}_p^{(k)}(\xi) \otimes \omega)^H \mathbf{V}_\gamma \mathbf{V}_\gamma^H (\mathbb{B}_p^{(k)}(\xi) \otimes \omega) = \sum_{p=1}^{N_k} \omega^H (\mathbb{B}_p^{(k)}(\xi) \otimes \mathbf{I}_M)^H \mathbf{V}_\gamma \mathbf{V}_\gamma^H (\mathbb{B}_p^{(k)}(\xi) \otimes \mathbf{I}_M) \omega$$
$$= \omega^H \Pi^{(k)}(\xi) \omega$$

where

$$\prod^{(k)}(\xi) = \sum_{p=1}^{N_k} (\mathbb{B}_p^{(k)}(\xi) \otimes \mathbf{I}_M)^H \mathbf{V}_\gamma \mathbf{V}_\gamma^H (\mathbb{B}_p^{(k)}(\xi) \otimes \mathbf{I}_M). \tag{19}$$

Lemma 1 tells us that when $\mathbf{A}$ has full column rank,
(1) the matrix $\Pi^{(k)}(\xi)$ is singular at $\xi = \xi^{(k)}$. Meanwhile, $\Pi^{(k)}(\xi^{(k)})$ has N_{cl} zero eigenvalues;
(2) the matrix $\Pi^{(k)}(\xi)$ should be positive definite when $\xi \neq \xi^{(k)}$.
It implies the matrix $\Pi^{(k)}(\xi)$ drops rank if and only if $\xi = \xi^{(k)}$. Based on above observations, we design the

CFO estimation as follows. Select a trial ξ from (-0.5, 0.5) and compute the M eigenvalues of the matrix $\Pi^{(k)}(\xi)$, denoted by $\kappa_1^{(k)}(\xi), \kappa_2^{(k)}(\xi), \dots, \kappa_M^{(k)}(\xi)$ in ascending order. The CFO for the kth user can be obtained from 1D search by minimizing the following cost function:

$$\hat{\xi}^{(k)} = \arg \min_\xi \sum_{l=1}^{N_{cl}} \kappa_l^{(k)}(\xi). \tag{20}$$

DOA estimation
We denote $\varepsilon_l^{(k)}(\xi)$ as the eigenvector of matrix $\Pi^{(k)}(\xi)$ corresponding to its lth eigenvalue $\kappa_l^{(k)}(\xi)$. Notice that the first N_{cl} eigenvectors $\varepsilon_l^{(k)}(\xi^{(k)})$, $l = 1, 2, \dots, N_{cl}$, correspond to the N_{cl} zero eigenvalues of $\Pi^{(k)}(\xi^{(k)})$. From Lemma 1, the N_{cl} column vectors of Vandermonde matrix $a^{(k)}$ constitute the same column space of $[\varepsilon_1^{(k)}(\xi^{(k)}), \varepsilon_2^{(k)}(\xi^{(k)}), \dots, \varepsilon_{N_{cl}}^{(k)}(\xi^{(k)})]$, which implies that $a^{(k)}$ should be orthogonal to the other $M - N_{cl}$ eigenvectors, i.e.,

$$(a^{(k)})^H \varepsilon_l^{(k)}(\xi^{(k)}) = \mathbf{0}, \quad l = N_{cl} + 1, \dots, M. \tag{21}$$

Thereby, after the CFO estimation for the kth user, we can further derive the N_{cl} DOAs for the kth user by finding the N_{cl} minimum point of the following cost function:

$$\hat{\varphi}_i^{(k)} = \arg \min_\varphi \sum_{l=N_{cl}+1}^{M} \left\| (\alpha(\varphi))^H \varepsilon_l^{(k)}(\hat{\xi}^{(k)}) \right\|^2 = \arg \min_\varphi g^{(k)}(\varphi), \tag{22}$$

$i = 1, 2, \dots, N_{cl}$, where $\alpha(\varphi) = [1, e^{j2\pi \chi \cos \varphi}, \dots, e^{j2\pi \chi(M-1)\cos \varphi}]^T$. Note that the polynomial rooting approach can be used to implement this minimization problem. The basic idea is first obtaining all local minimum/maximum solutions by setting the derivative of the cost function to be zero, and then putting these solutions back to the original cost function and selecting the minimum after comparison [22]. Specifically, denoting $\Theta^{(k)}(\xi) = \sum_{l=N_{cl}+1}^{M} \varepsilon_l^{(k)}(\xi)(\varepsilon_l^{(k)}(\xi))^H$ and $\Psi = \mathrm{diag}(0, 1, \dots, M-1)$, we obtain

$$\frac{\partial g^{(k)}(\varphi)}{\partial \varphi} = j2\pi \chi \sin \varphi \cdot (\alpha(\varphi))^H (\Psi \Theta^{(k)}(\hat{\xi}^{(k)}) - \Theta^{(k)}(\hat{\xi}^{(k)})\Psi) \alpha(\varphi). \tag{23}$$

According to [22], we know $z = e^{j2\pi \chi \cos \hat{\varphi}_i^{(k)}}$, $i = 1, 2, \dots, N_{cl}$, is one of the roots for the polynomial $Q(z) = \sum_{m=-M+1}^{M-1} b_m z^m$ where

$$b_m = \sum_{q-p=m} [\Psi \Theta^{(k)}(\hat{\xi}^{(k)}) - \Theta^{(k)}(\hat{\xi}^{(k)})\Psi]_{pq}. \tag{24}$$

By putting the roots on the unit circle of $Q(z)$ back to the original cost function $g^{(k)}(\phi)$ and selecting the N_{cl} minimum points after comparisons, the solution to (22) is obtained. Afterwards, for $\underline{A}$ with full column rank, we obtain Lemma 2:

Lemma 2: Assume all the K users have distinct DOAs. According to the number of subcarriers allocated, we arrange the user in descending order, as follows

$$e_1, e_2, \ldots, e_K,$$

such that $N_{e_k} \geq N_{e_{k+1}}$, then when

$$\begin{cases} L \geq N_{e_1} + 1, & v \geq K, \\ L \geq N_{e_1} + 1 + \sum_{k=v+1}^{K} N_{e_k}, & v < K, \end{cases} \quad (25)$$

where $v = \left\lfloor \frac{M}{N_{cl}} \right\rfloor \geq 2$ with $\lfloor \cdot \rfloor$ denoting the integer floor operation, the matrix $\underline{A}$ has full column rank.

Proof. See Appendix 2. □

Note that when $\underline{A}$ has full column rank, the matrix $\mathbf{A}$ will be tall and also has full column rank, which guarantees the validity of the SVD operation in (15). Following Lemmas 1 and 2, we could make the following important observation; that is when $M \geq 2N_{cl}$, i.e., the number of antennas at the receivers is not less than two times of the number of channel clusters from each user, the validity of our method is guaranteed with L satisfying the condition in (25).

Performance analysis

In this section, we provide theoretical performance analysis for both the CFO and DOA estimation performance of our proposed method. Bearing in mind that $\mathbf{V}_\gamma \mathbf{V}_\gamma^H = I_{ML} - U_\gamma U_\gamma^H$,

we can rewrite $\mathbf{\Pi}^{(k)}(\zeta)$ as follows

$$\mathbf{\Pi}^{(k)}(\xi) = \sum_{p=1}^{N_k} \left(\frac{L}{N} \mathbf{I}_M - (\mathbb{B}_p^{(k)}(\xi) \otimes \mathbf{I}_M)^H \mathbf{U}_\gamma \mathbf{U}_\gamma^H (\mathbb{B}_p^{(k)}(\xi) \otimes \mathbf{I}_M) \right)$$
$$= \frac{N_k L}{N} \mathbf{I}_M - \mathbf{\Xi}^{(k)}(\xi) \quad (26)$$

where

$$\mathbf{\Xi}^{(k)}(\xi) = \sum_{p=1}^{N_k} (\mathbb{B}_p^{(k)}(\xi) \otimes \mathbf{I}_M)^H \mathbf{U}_\gamma \mathbf{U}_\gamma^H (\mathbb{B}_p^{(k)}(\xi) \otimes \mathbf{I}_M). \quad (27)$$

Denote the M eigenvalues of matrix $\mathbf{\Xi}^{(k)}(\zeta)$ by $\lambda_1^{(k)}(\xi)$, $\lambda_2^{(k)}(\xi), \ldots, \lambda_M^{(k)}(\xi)$ in ascending order. The corresponding eigenvectors are denoted by $\boldsymbol{v}_1^{(k)}(\xi)$, $\boldsymbol{v}_2^{(k)}(\xi), \ldots, \boldsymbol{v}_M^{(k)}(\xi)$. We can readily verify the following relationships: $\lambda_l^{(k)}(\xi) = \frac{N_k L}{N} - \kappa_{M-l+1}^{(k)}(\xi)$ and $\boldsymbol{v}_l^{(k)}(\xi) = \boldsymbol{\varepsilon}_{M-l+1}^{(k)}(\xi)$.

Then, (20) and (22) can be rewritten as follows:

$$\hat{\xi}^{(k)} = \arg \max_\xi \sum_{l=M-N_{c1}+1}^{M} \lambda_l^{(k)}(\xi) = \arg \max_\xi G^{(k)}(\xi), (28)$$

$$\hat{\varphi}_i^{(k)} = \arg \min_\varphi \sum_{l=1}^{M-N_{cl}} \left\| (\boldsymbol{\alpha}(\varphi))^H \boldsymbol{v}_l^{(k)}(\hat{\xi}^{(k)}) \right\|^2 = \arg \min_\varphi D^{(k)}(\varphi). \quad (29)$$

CFO estimation performance

We rewrite (15) as

$$\mathbf{R}_\gamma = \sum_{i=1}^{ML} \sigma_i^2 e_i e_i^H \quad (30)$$

where σ_i^2 and e_i, $i = 1, 2, \ldots, N_{cl}N_{sum}$, denote the eigenvalues and eigenvectors corresponding to signal space, respectively, while the remaining $\sigma_i^2 = \sigma_n^2$ and e_i, $i = N_{cl} N_{sum} + 1, \ldots, ML$, correspond to the noise space.

Let $\boldsymbol{\eta}_i^\Re = \Re\{e_i\}$ and $\boldsymbol{\eta}_i^\Im = \Im\{e_i\}$, i.e., $e_i = \boldsymbol{\eta}_i^\Re + j\boldsymbol{\eta}_i^\Im$. Denote $\eta_i = [(\boldsymbol{\eta}_i^\Re)^T, (\boldsymbol{\eta}_i^\Im)^T]^T$, and $\boldsymbol{\eta} = [\boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T, \ldots, \boldsymbol{\eta}_{N_{cl} N_{sum}}^T]^T$. Let $\hat{x}$ stand for the estimated value for x. The CFO estimation variance of the kth user can be given by [23]:

$$Cov\{\hat{\xi}^{(k)}\} = \left[\frac{\partial^2 G^{(k)}}{\partial \xi^2} \right]^{-1} \left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \boldsymbol{\eta}} \right]$$
$$Cov\{\hat{\boldsymbol{\eta}}\} \left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \boldsymbol{\eta}} \right]^T \left[\frac{\partial^2 G^{(k)}}{\partial \xi^2} \right]^{-1} \Bigg|_{\xi = \xi^{(k)}}. \quad (31)$$

We denote $\mathbf{\Gamma}_p^{(k)}(\xi) = \mathbb{B}_p^{(k)}(\xi) \otimes \mathbf{I}_M$ for short. Then, there is

$$\frac{\partial \mathbf{\Gamma}_p^{(k)}(\xi)}{\partial \xi} = D \cdot \mathbf{\Gamma}_p^{(k)}(\xi) \quad (32)$$

Where $D = \frac{j2\pi}{N} \text{diag}(0, 1, \ldots, L-1) \otimes \mathbf{I}_M$. In the following, we omit the parameterized notation (ξ) for presentation clarity. Then, we have

$$\frac{\partial G^{(k)}}{\partial \boldsymbol{\xi}} = \sum_{l=M-N_{cl}+1}^{M} (\boldsymbol{v}_l^{(k)})^H \frac{\partial \mathbf{\Xi}^{(k)}}{\partial \boldsymbol{\xi}} \boldsymbol{v}_l, \quad (33)$$

$$\frac{\partial^2 G^{(k)}}{\partial \xi^2} = \sum_{l=M-N_{cl}+1}^{M} \left[(\boldsymbol{v}_l^{(k)})^H \frac{\partial^2 \mathbf{\Xi}^{(k)}}{\partial \xi^2} \boldsymbol{v}_l^{(k)} \right.$$
$$\left. + 2\Re \left\{ (\boldsymbol{v}_l^{(k)})^H \frac{\partial \mathbf{\Xi}^{(k)}}{\partial \xi} \left(\frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \xi} \right) \right\} \right] \quad (34)$$

where

$$\frac{\partial \Xi^{(k)}}{\partial \xi} = \sum_{p=1}^{N_k} \left((\Gamma_p^{(k)})^H D^H \mathbf{U}_\gamma \mathbf{U}_\gamma^H \Gamma_p^{(k)} + (\Gamma_p^{(k)})^H \mathbf{U}_\gamma \mathbf{U}_\gamma^H D \Gamma_p^{(k)} \right), \quad (35)$$

$$\frac{\partial^2 \Xi^{(k)}}{\partial \xi^2} = \sum_{p=1}^{N_k} \left((\Gamma_p^{(k)})^H D^H D^H \mathbf{U}_\gamma \mathbf{U}_\gamma^H \Gamma_p^{(k)} + 2(\Gamma_p^{(k)})^H D^H \mathbf{U}_\gamma \mathbf{U}_\gamma^H D \Gamma_p^{(k)} \right.$$
$$\left. + (\Gamma_p^{(k)})^H \mathbf{U}_\gamma \mathbf{U}_\gamma^H D D \Gamma_p^{(k)} \right), \quad (36)$$

$$\frac{\partial \mathbf{v}_l^{(k)}}{\partial \xi} = \sum_{z \neq l}^{M} \frac{(\mathbf{v}_z^{(k)})^H \frac{\partial \Xi^{(k)}}{\partial \xi} \mathbf{v}_l^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}} \mathbf{v}_z^{(k)}. \quad (37)$$

Next, we obtain

$$\frac{\partial^2 G^{(k)}}{\partial \xi \partial \eta_i^\Re} = \sum_{l=M-N_{cl}+1}^{M} (\mathbf{v}_l^{(k)})^H \frac{\partial 2 \Xi^{(k)}}{\partial \xi \partial \eta_i^\Re} \mathbf{v}_l^{(k)}$$
$$+ 2\Re \left\{ \sum_{l=M-N_{cl}+1}^{M} (\mathbf{v}_l^{(k)})^H \frac{\partial \Xi^{(k)}}{\partial \xi} \frac{\partial \mathbf{v}_l^{(k)}}{\partial \eta_i^\Re} \right\} \quad (38)$$

where

$$\frac{\partial^2 \Xi^{(k)}}{\partial \xi \partial \eta_i^\Re} = \sum_{p=1}^{N_k} \left((\Gamma_p^{(k)})^H D^H \frac{\partial (\mathbf{U}_\gamma \mathbf{U}_\gamma^H)}{\partial \eta_i^\Re} \Gamma_p^{(k)} \right.$$
$$\left. + (\Gamma_p^{(k)})^H \frac{\partial (\mathbf{U}_\gamma \mathbf{U}_\gamma^H)}{\partial \eta_i^\Re} D \Gamma_p^{(k)} \right), \quad (39)$$

$$\frac{\partial \mathbf{v}_l^{(k)}}{\partial \eta_i^\Re} = \sum_{z \neq l}^{M} \frac{\mathbf{v}_z^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}} \left((\mathbf{v}_z^{(k)})^H \frac{\partial \Xi^{(k)}}{\partial \eta_i^\Re} \mathbf{v}_l^{(k)} \right)$$
$$= \sum_{z \neq l}^{M} \frac{\mathbf{v}_z^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}} \quad (40)$$
$$\left(\sum_{p=1}^{N_k} (\mathbf{v}_z^{(k)})^H (\Gamma_p^{(k)})^H \frac{\partial (\mathbf{U}_\gamma \mathbf{U}_\gamma^H)}{\partial \eta_i^\Re} \Gamma_p^{(k)} \mathbf{v}_l^{(k)} \right).$$

Using

$$\omega_1^H \frac{\partial (\mathbf{U}_\gamma \mathbf{U}_\gamma^H)}{\partial \eta_i^\Re} \omega_2 = e_i^H \omega_2 \omega_1^H + \omega_1^H e_i \omega_2^T \quad (41)$$

and skipping some algebraic steps, we can simplify (38) as

$$\frac{\partial^2 G^{(k)}}{\partial \xi \partial \eta_i^\Re} = 2\Re\{e_i^H \mathbf{Q}^{(k)}\} \quad (42)$$

where

$$\mathbf{Q}^{(k)} = \tilde{\mathbf{Q}}^{(k)} + (\tilde{\mathbf{Q}}^{(k)})^H, \quad (43)$$

$$\tilde{\mathbf{Q}}^{(k)} = \sum_{l=M-N_{cl}+1}^{M} \sum_{p=1}^{N_k} \left(\Gamma_p^{(k)} \mathbf{v}_l^{(k)} (\mathbf{v}_l^{(k)})^H (\Gamma_p^{(k)})^H D^H \right.$$
$$+ \sum_{z=1}^{M-N_{cl}} \frac{(\mathbf{v}_l^{(k)})^H (\frac{\partial \Xi^{(k)}}{\partial \xi}) \mathbf{v}_z^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}} \quad (44)$$
$$\left. \Gamma_p^{(k)} \mathbf{v}_l^{(k)} (\mathbf{v}_z^{(k)})^H (\Gamma_p^{(k)})^H \right).$$

Likewise, we can also obtain

$$\frac{\partial^2 G^{(k)}}{\partial \xi \partial \eta_i^\Im} = -2\Im\{e_i^H \mathbf{Q}^{(k)}\}. \quad (45)$$

Furthermore, using

$$\text{Cov}\{\hat{\eta}_i, \hat{\eta}_j\} = \frac{1}{2} \begin{bmatrix} \Re\{\text{Cov}\{\hat{e}_i, \hat{e}_j\} + \text{Cov}\{\hat{e}_i, \hat{e}_j^*\}\} & \Im\{-\text{Cov}\{\hat{e}_i, \hat{e}_j\} + \text{Cov}\{\hat{e}_i, \hat{e}_j^*\}\} \\ \Im\{\text{Cov}\{\hat{e}_i, \hat{e}_j\} + \text{Cov}\{\hat{e}_i, \hat{e}_j^*\}\} & \Re\{\text{Cov}\{\hat{e}_i, \hat{e}_j\} - \text{Cov}\{\hat{e}_i, \hat{e}_j^*\}\} \end{bmatrix}, \quad (46)$$

we can further obtain

$$\frac{\partial^2 G^{(k)}}{\partial \xi \partial \eta} \text{Cov}\{\hat{\eta}\} \left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \eta} \right]^T$$
$$= 2\Re \left\{ \sum_{i=1}^{N_{cl}N_{sum}} \sum_{j=1}^{N_{cl}N_{sum}} \left(e_i^H \mathbf{Q}^{(k)} \text{Cov}\{\hat{e}_i, \hat{e}_j\} \mathbf{Q}^{(k)} e_j \right. \right. \quad (47)$$
$$\left. \left. + e_i^H \mathbf{Q}^{(k)} \text{Cov}\{\hat{e}_i, \hat{e}_j^*\} (\mathbf{Q}^{(k)})^* e_j^* \right) \right\}.$$

Based on the results from [24], we know

$$\text{Cov}\{\hat{e}_i, \hat{e}_j\} = \sum_{q=1,q\neq i}^{ML} \sum_{f=1,f\neq j}^{ML} \left[\frac{1}{L_s(N-L+1)^2} \right.$$
$$\left. \sum_{t=0}^{N-L} \sum_{r=0}^{N-L} \frac{e_q^H \mathbf{R}_{t,r} e_f e_j^H \mathbf{R}_{r,t} e_i}{(\sigma_i^2 - \sigma_q^2)(\sigma_j^2 - \sigma_f^2)} \right] e_q e_f^H, \quad (48)$$

$$\text{Cov}\{\hat{e}_i, \hat{e}_j^*\} = \sum_{q=1,q\neq i}^{ML} \sum_{f=1,f\neq j}^{ML} \left[\frac{1}{L_s(N-L+1)^2} \right.$$
$$\left. \sum_{t=0}^{N-L} \sum_{r=0}^{N-L} \frac{e_q^H \mathbf{R}_{t,r} e_j e_f^H \mathbf{R}_{r,t} e_i}{(\sigma_i^2 - \sigma_q^2)(\sigma_j^2 - \sigma_f^2)} \right] e_q e_f^T \quad (49)$$

where

$$\mathbf{R}_{t,r} = E[\gamma_g|_t^{t+L-1} (\gamma_g|_r^{r+L-1})^H] = \sigma_s^2 \mathbf{A} \mathbf{\Phi}^{t-r} \mathbf{A}^H + \sigma_n^2 \mathbf{J}_{t,r}, \quad (50)$$

and $\mathbf{J}_{tr}$ is a submatrix of $\mathbf{I}_{N_{cl}N_{sum}}$ which is given by $\mathbf{J}_{t,r} = \mathbf{I}_{N_{cl}N_{sum}} (1 + (t-1)M : 1 + (t+L-2)M, 1+(r-1)M : 1+(r+L-2)M)$.

Substituting (48) and (49) into (47) and after some algebraic steps, we arrive at

$$
\left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \boldsymbol{\eta}}\right] \mathrm{Cov}\{\hat{\boldsymbol{\eta}}\} \left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \boldsymbol{\eta}}\right]^T
$$

$$
= \frac{2}{L_s(N-L+1)^2} \Re \left[\sum_{i=1}^{N_{\mathrm{cl}}} N_{\mathrm{sum}} \sum_{j=1}^{N_{\mathrm{cl}} N_{\mathrm{sum}}} \sum_{q=N_{\mathrm{cl}} N_{\mathrm{sum}}+1}^{ML} \right.
$$

$$
\sum_{f=N_{\mathrm{cl}} N_{\mathrm{sum}}+1}^{ML} \sum_{t=0}^{N-L} \sum_{r=0}^{N-L} \left(e_i^H Q^{(k)} \frac{e_q^H \mathbf{R}_{t,r} e_f e_j^H \mathbf{R}_{r,t} e_i}{(\sigma_i^2 - \sigma_n^2)(\sigma_j^2 - \sigma_n^2)} \right. \tag{51}
$$

$$
e_q e_f^H \mathbf{Q}^{(k)} e_j + e_i^H \mathbf{Q}^{(k)}
$$

$$
\left. \left. \frac{e_q^H \mathbf{R}_{t,r} e_j e_f^H \mathbf{R}_{r,t} e_i}{(\sigma_i^2 - \sigma_n^2)(\sigma_j^2 - \sigma_n^2)} e_q e_f^T (\mathbf{Q}^{(k)})^* (e_j)^* \right) \right].
$$

By defining

$$
\boldsymbol{\Omega}_{UV}^{(k)} = \mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{Q}^{(k)} \mathbf{V}_{\boldsymbol{\gamma}}, \quad \boldsymbol{\Omega}_{VU}^{(k)} = \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{Q}^{(k)} \mathbf{U}_{\boldsymbol{\gamma}}, \quad \boldsymbol{\Delta}_{t,r,UU}^{(k)} = \mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{R}_{t,r} \mathbf{U}_{\boldsymbol{\gamma}},
$$

$$
\boldsymbol{\Delta}_{t,r,VV}^{(k)} = \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{R}_{t,r} \mathbf{V}_{\boldsymbol{\gamma}}, \quad \boldsymbol{\Delta}_{t,r,UV}^{(k)} = \mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{R}_{t,r} \mathbf{V}_{\boldsymbol{\gamma}}, \quad \boldsymbol{\Delta}_{t,r,VU}^{(k)} = \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{R}_{t,r} \mathbf{U}_{\boldsymbol{\gamma}}
$$

we can further rewrite (51) into the following more compact form

$$
\left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \boldsymbol{\eta}}\right] \mathrm{Cov}\{\hat{\boldsymbol{\eta}}\} \left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \boldsymbol{\eta}}\right]^T
$$

$$
= \frac{2}{L_s(N-L+1)^2} \Re \left[\sum_{t=0}^{N-L} \sum_{r=0}^{N-L} d_\sigma^T \left(\boldsymbol{\Omega}_{UV}^{(k)} \boldsymbol{\Delta}_{t,r,VV}^{(k)} \boldsymbol{\Omega}_{VU}^{(k)} \right) \circ (\boldsymbol{\Delta}_{t,r,UU}^{(k)})^* \right. \tag{52}
$$

$$
\left. + (\boldsymbol{\Omega}_{UV}^{(k)} \boldsymbol{\Delta}_{t,r,VU}^{(k)}) \circ ((\boldsymbol{\Delta}_{t,r,UV}^{(k)})^* (\boldsymbol{\Omega}_{UV}^{(k)})^T) \right) d_\sigma \right]
$$

where $d_\sigma = \left[\frac{1}{\sigma_1^2 - \sigma_n^2}, \frac{1}{\sigma_2^2 - \sigma_n^2}, \cdots, \frac{1}{\sigma_{N_{\mathrm{cl}} N_{\mathrm{sum}}}^2 - \sigma_n^2}\right]^T$.

Based on the fact that $\mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{A} = \mathbf{0}$, we further obtain

$$
\boldsymbol{\Delta}_{t,r,UU}^{(k)} = \sigma_s^2 \mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{A} \boldsymbol{\Phi}^{t-r} \mathbf{A}^H \mathbf{U}_{\boldsymbol{\gamma}} + \sigma_n^2 \mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{J}_{t,r} \mathbf{U}_{\boldsymbol{\gamma}} \simeq \sigma_s^2 \mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{A} \boldsymbol{\Phi}^{t-r} \mathbf{A}^H \mathbf{U}_{\boldsymbol{\gamma}},
$$

$$
\boldsymbol{\Delta}_{t,r,VV}^{(k)} = \sigma_n^2 \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{J}_{t,r} \mathbf{V}_{\boldsymbol{\gamma}}, \quad \boldsymbol{\Delta}_{t,r,UV}^{(k)} = \sigma_n^2 \mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{J}_{t,r} \mathbf{V}_{\boldsymbol{\gamma}}, \quad \boldsymbol{\Delta}_{t,r,VU}^{(k)} = \sigma_n^2 \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{J}_{t,r} \mathbf{U}_{\boldsymbol{\gamma}}.
$$

Then, we have

$$
(\boldsymbol{\Omega}_{UV}^{(k)} \boldsymbol{\Delta}_{t,r,VV}^{(k)} \boldsymbol{\Omega}_{VU}^{(k)}) \, (\boldsymbol{\Delta}_{t,r,UU}^{(k)})^* = \sigma_s^2 \sigma_n^2 (\boldsymbol{\Omega}_{UV}^{(k)} \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{J}_{t,r} \mathbf{V}_{\boldsymbol{\gamma}} \boldsymbol{\Omega}_{VU}^{(k)}) \, (\mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{A} \boldsymbol{\Phi}^{t-r} \mathbf{A}^H \mathbf{U}_{\boldsymbol{\gamma}})^*, \tag{53}
$$

$$
(\boldsymbol{\Omega}_{UV}^{(k)} \boldsymbol{\Delta}_{t,r,VU}^{(k)}) \, ((\boldsymbol{\Delta}_{t,r,UV}^{(k)})^* (\boldsymbol{\Omega}_{UV}^{(k)})^T) = \sigma_n^4 (\boldsymbol{\Omega}_{UV}^{(k)} \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{J}_{t,r} \mathbf{U}_{\boldsymbol{\gamma}}) \, ((\mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{J}_{t,r} \mathbf{V}_{\boldsymbol{\gamma}})^* (\boldsymbol{\Omega}_{UV}^{(k)})^T). \tag{54}
$$

Bearing in mind that (54) is infinitesimal under high signal-to-noise ratio (SNR) as compared to (53), we can then approximate (52) as

$$
\left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \boldsymbol{\eta}}\right] \mathrm{Cov}\{\hat{\boldsymbol{\eta}}\} \left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \boldsymbol{\eta}}\right]^T
$$

$$
\simeq \frac{2\sigma_s^2 \sigma_n^2}{L_s(N-L+1)^2} \Re \left[\sum_{t=0}^{N-L} \sum_{r=0}^{N-L} d_\sigma^T ((\boldsymbol{\Omega}_{UV}^{(k)} \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{J}_{t,r} \mathbf{V}_{\boldsymbol{\gamma}} \boldsymbol{\Omega}_{VU}^{(k)}) \, (\mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{A} \boldsymbol{\Phi}^{t-r} \mathbf{A}^H \mathbf{U}_{\boldsymbol{\gamma}})^*) d_\sigma \right]. \tag{55}
$$

Let $\tilde{\sigma}_i^2$, $i = 1, 2, \ldots, ML$, be the eigenvalues of matrix $\mathbf{A} \mathbf{A}^H$. Then, there holds $\sigma_i^2 = \sigma_s^2 \tilde{\sigma}_i^2 + \sigma_n^2$. By defining $\tilde{d}_\sigma = \left[\frac{1}{\tilde{\sigma}_1^2}, \frac{1}{\tilde{\sigma}_2^2}, \cdots, \frac{1}{\tilde{\sigma}_{N_{\mathrm{cl}} N_{\mathrm{sum}}}^2}\right]^T$, we have $\tilde{d}_\sigma = d_\sigma / \sigma_s^2$. Then, we can rewrite (55) as

$$
\left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \boldsymbol{\eta}}\right] \mathrm{Cov}\{\hat{\boldsymbol{\eta}}\} \left[\frac{\partial^2 G^{(k)}}{\partial \xi \partial \boldsymbol{\eta}}\right]^T
$$

$$
\simeq \frac{2\sigma_n^2}{\sigma_s^2 L_s(N-L+1)^2} \tag{56}
$$

$$
\Re \left[\sum_{t=0}^{N-L} \sum_{r=0}^{N-L} \tilde{d}_\sigma^T ((\boldsymbol{\Omega}_{UV}^{(k)} \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{J}_{t,r} \mathbf{V}_{\boldsymbol{\gamma}} \boldsymbol{\Omega}_{VU}^{(k)}) \circ (\mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{A} \boldsymbol{\Phi}^{t-r} \mathbf{A}^H \mathbf{U}_{\boldsymbol{\gamma}})^*) \tilde{d}_\sigma \right].
$$

On the other side, we can also rewrite $\frac{\partial^2 G^{(k)}}{\partial \xi^2}$ as follows

$$
\left. \frac{\partial^2 G^{(k)}}{\partial \xi^2} \right|_{\xi = \xi^{(k)}}
$$

$$
= 2 \sum_{l=M-N_{\mathrm{cl}}+1}^{M} \left[-(\boldsymbol{v}_l^{(k)})^H \boldsymbol{\Lambda}_1^{(k)} \boldsymbol{v}_l^{(k)} + \sum_{z=1}^{M-N_{\mathrm{cl}}} \frac{|(\boldsymbol{v}_l^{(k)})^H \boldsymbol{\Lambda}_2^{(k)} \boldsymbol{v}_z^{(k)}|^2}{\lambda_l^{(k)} - \lambda_z^{(k)}} \right], \tag{57}
$$

where

$$
\boldsymbol{\Lambda}_1^{(k)} = \sum_{p=1}^{N_k} (\boldsymbol{\Gamma}_p^{(k)})^H \mathbf{D}^H \mathbf{V}_{\boldsymbol{\gamma}} \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{D} \boldsymbol{\Gamma}_p^{(k)}, \quad \boldsymbol{\Lambda}_2^{(k)} = \sum_{p=1}^{N_k} (\boldsymbol{\Gamma}_p^{(k)})^H \mathbf{D}^H \mathbf{V}_{\boldsymbol{\gamma}} \mathbf{V}_{\boldsymbol{\gamma}}^H \boldsymbol{\Gamma}_p^{(k)}.
$$

Finally, substituting (55) and (57) into (31), we arrive at

$$
\mathrm{Cov}\{\hat{\xi}^{(k)}\}
$$

$$
\simeq \frac{\sigma_n^2}{2\sigma_s^2 L_s(N-L+1)^2}
$$

$$
\cdot \left. \frac{\Re \left[\sum_{t=0}^{N-L} \sum_{r=0}^{N-L} \tilde{d}_\sigma^T ((\boldsymbol{\Omega}_{UV}^{(k)} \mathbf{V}_{\boldsymbol{\gamma}}^H \mathbf{J}_{t,r} \mathbf{V}_{\boldsymbol{\gamma}} \boldsymbol{\Omega}_{VU}^{(k)}) \circ (\mathbf{U}_{\boldsymbol{\gamma}}^H \mathbf{A} \boldsymbol{\Phi}^{t-r} \mathbf{A}^H \mathbf{U}_{\boldsymbol{\gamma}})^*) \tilde{d}_\sigma \right]}{\left[\sum_{l=M-N_{\mathrm{cl}}+1}^{M} -(\boldsymbol{v}_l^{(k)})^H \boldsymbol{\Lambda}_1^{(k)} \boldsymbol{v}_l^{(k)} + \sum_{z=1}^{M-N_{\mathrm{cl}}} \frac{|(\boldsymbol{v}_l^{(k)})^H \boldsymbol{\Lambda}_2^{(k)} \boldsymbol{v}_z^{(k)}|^2}{\lambda_l^{(k)} - \lambda_z^{(k)}} \right]^2} \right|_{\xi = \xi^{(k)}} \tag{58}
$$

From (58), we make the following observations: First, it is seen that, the CFO estimation variance is decreased when SNR, i.e., σ_s^2 / σ_n^2, increases. Second, increasing the number of blocks, i.e., L_s, also improves the CFO estimation performance. Third, the relationship between the value of L and the CFO estimation variance is quite complicated. We see that, decreasing the value of L, on the one hand, will decrease the value of $\frac{1}{(N-L+1)^2}$; on the other hand, it will also decrease the eigenvalues $\tilde{\sigma}_i^2$, resulting in larger entries of $\tilde{d}_\sigma$. In fact, from (11) and (14), we see that, the effect of L resembles the smoothing technique in array signal processing. That is, decreasing the value of L, on the one hand, will reduce the fluctuations of the estimated correlation matrix $\hat{R}_{\boldsymbol{\Gamma}}$; on the other hand, it also decreases the potential for higher resolution of the subspace algorithm. We will later investigate the impact of L on the estimation performance via numerical simulations.

DOA estimation performance

Denote $\boldsymbol{\mu}_l^{(k)} = [\Re\{(\boldsymbol{v}_l^{(k)})^T\}, \ \Im\{(\boldsymbol{v}_l^{(k)})^T\}]^T$, and $\boldsymbol{\mu}^{(k)} = [(\boldsymbol{\mu}_1^{(k)})^T, (\boldsymbol{\mu}_2^{(k)})^T, \ldots, (\boldsymbol{\mu}_{M-N_{c1}}^{(k)})^T]^T$. Likewise, the DOA estimation variance of the kth user is given by

$$\mathrm{Cov}\{\hat{\varphi}_i^{(k)}\} = \left[\frac{\partial^2 D^{(k)}}{\partial \varphi^2}\right]^{-1} \left[\frac{\partial^2 D^{(k)}}{\partial \varphi \partial \boldsymbol{\mu}^{(k)}}\right]$$
$$Cov\{\hat{\boldsymbol{\mu}}^{(k)}\} \left[\frac{\partial^2 D^{(k)}}{\partial \varphi \partial \boldsymbol{\mu}^{(k)}}\right]^T \left[\frac{\partial^2 D^{(k)}}{\partial \varphi^2}\right]^{-1} \Bigg|_{\varphi = \varphi_i^{(k)}}. \tag{59}$$

We have

$$\frac{\partial D^{(k)}}{\partial \varphi} = 2\Re\left\{(\boldsymbol{\alpha}(\varphi))^H \left(\boldsymbol{v}^{(k)}(\boldsymbol{v}^{(k)})^H\right)\boldsymbol{\Psi}_1\boldsymbol{\alpha}(\varphi)\right\}, \tag{60}$$

$$\frac{\partial^2 D^{(k)}}{\partial \varphi^2} = 2\Re\left\{(\boldsymbol{\alpha}(\varphi))^H \boldsymbol{\Psi}_1^H \left(\boldsymbol{v}^{(k)}(\boldsymbol{v}^{(k)})^H\right)\boldsymbol{\Psi}_1\boldsymbol{\alpha}(\varphi) \right.$$
$$\left. + (\boldsymbol{\alpha}(\varphi))^H \left(\boldsymbol{v}^{(k)}(\boldsymbol{v}^{(k)})^H\right)\boldsymbol{\Psi}_1(\boldsymbol{\Psi}_1 + \boldsymbol{\Psi}_2)\boldsymbol{\alpha}(\varphi)\right\} \tag{61}$$

where

$\boldsymbol{v}^{(k)} = \left[\boldsymbol{v}_1^{(k)}, \ \boldsymbol{v}_2^{(k)}, \ \ldots, \ \boldsymbol{v}_{M-N_{c1}}^{(k)}\right]$, $\boldsymbol{\Psi}_1 = -\mathrm{j}2\pi\chi\sin(\varphi)\mathrm{diag}(0, 1, \ldots, M-1)$, $\boldsymbol{\Psi}_2 = -\mathrm{j}2\pi\chi\cos(\varphi)\mathrm{diag}(0, 1, \ldots, M-1)$.

We further obtain

$$\left[\frac{\partial^2 D}{\partial \varphi \partial \boldsymbol{\mu}^{(k)}}\right]\mathrm{Cov}\{\hat{\boldsymbol{\mu}}^{(k)}\}\left[\frac{\partial^2 D}{\partial \varphi \partial \boldsymbol{\mu}^{(k)}}\right]^T$$
$$= \sum_{l=1}^{M-N_{c1}}\sum_{z=1}^{M-N_{c1}} 2\Re\left\{(\boldsymbol{v}_l^{(k)})^H\mathbf{T}^{(k)}\mathrm{Cov}\{\hat{\boldsymbol{v}}_l^{(k)}, \hat{\boldsymbol{v}}_z^{(k)}\}\mathbf{T}^{(k)}\boldsymbol{v}_z^{(k)} \right. \tag{62}$$
$$\left. (\boldsymbol{v}_l^{(k)})^H\mathbf{T}^{(k)}Cov\{\hat{\boldsymbol{v}}_l^{(k)}, (\hat{\boldsymbol{v}}_z^{(k)})^*\}(\mathbf{T}^{(k)})^*(\hat{\boldsymbol{v}}_z^{(k)})^*\right\}$$

where

$$\mathbf{T}^{(k)} = \boldsymbol{\Psi}_1\boldsymbol{\alpha}(\varphi)(\boldsymbol{\alpha}(\varphi))^H + \boldsymbol{\alpha}(\varphi)(\boldsymbol{\alpha}(\varphi))^H\boldsymbol{\Psi}_1^H. \tag{63}$$

Let Δx denote the perturbation of x, i.e., $\Delta x = \hat{x} - x$. Bearing in mind that $\boldsymbol{v}_l^{(k)}$, $l = 1, 2, \ldots, M$, denotes the eigenvector of matrix $\boldsymbol{\Xi}^{(k)}(\hat{\boldsymbol{\xi}}^{(k)})$, its perturbation is introduced by both the CFO estimation error of the kth user and the perturbations of the eigenvectors $\boldsymbol{e}_i$, $i = 1, 2, \ldots, N_{c1}N_{sum}$. Hence, using the first-order approximation, the perturbation of $\boldsymbol{v}_l^{(k)}$ can be expressed as

$$\Delta \boldsymbol{v}_l^{(k)} = \frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \xi}\Delta \xi^{(k)}$$
$$+ \sum_{i=1}^{N_{c1}N_{sum}}\left(\frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \boldsymbol{\eta}_i^\Re}\Delta \boldsymbol{\eta}_i^\Re + \frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \boldsymbol{\eta}_i^\Im}\Delta \boldsymbol{\eta}_i^\Im\right) = \boldsymbol{\Upsilon}_l^{(k)}\cdot\Delta\boldsymbol{\eta} \tag{64}$$

where

$$\Upsilon_l^{(k)} = -\frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \boldsymbol{\xi}}\left[\frac{\partial^2 G}{\partial \xi^2}\right]^{-1}\frac{\partial^2 G}{\partial \xi \partial \boldsymbol{\eta}} + \frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \boldsymbol{\eta}},$$

$$\frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \boldsymbol{\eta}_i^\Re} = \sum_{z\neq l}^M \frac{\boldsymbol{v}_l^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}}\sum_{p=1}^{N_k}\boldsymbol{e}_i^H\boldsymbol{\Gamma}_p^{(k)}\boldsymbol{v}_l^{(k)}(\boldsymbol{v}_z^{(k)})^H(\boldsymbol{\Gamma}_p^{(k)})^H + (\boldsymbol{v}_z^{(k)})^H(\boldsymbol{\Gamma}_p^{(k)})^H\boldsymbol{e}_i(\boldsymbol{v}_l^{(k)})^T(\boldsymbol{\Gamma}_p^{(k)})^T,$$

$$\frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \boldsymbol{\eta}_i^\Im} = \sum_{z\neq l}^M \mathrm{j}\frac{\boldsymbol{v}_z^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}}\sum_{p=1}^{N_k}\boldsymbol{e}_i^H\boldsymbol{\Gamma}_p^{(k)}\boldsymbol{v}_l^{(k)}(\boldsymbol{v}_z^{(k)})^H(\boldsymbol{\Gamma}_p^{(k)})^H - (\boldsymbol{v}_z^{(k)})^H(\boldsymbol{\Gamma}_p^{(k)})^H\boldsymbol{e}_i(\boldsymbol{v}_l^{(k)})^T(\boldsymbol{\Gamma}_p^{(k)})^T.$$

From

$$\mathrm{Cov}\{\hat{\boldsymbol{v}}_l^{(k)}, \hat{\boldsymbol{v}}_z^{(k)}\} = \boldsymbol{\Upsilon}_l^{(k)}Cov\{\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\eta}}\}(\boldsymbol{\Upsilon}_z^{(k)})^H \ \text{ and } \ \mathrm{Cov}\{\hat{\boldsymbol{v}}_l^{(k)}, (\hat{\boldsymbol{v}}_z^{(k)})^*\} = \boldsymbol{\Upsilon}_l^{(k)}Cov\{\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\eta}}\}(\boldsymbol{\Upsilon}_z^{(k)})^T,$$

we simplify (62) as follows after some manipulations:

$$\left[\frac{\partial^2 D}{\partial \varphi \partial \boldsymbol{\mu}^{(k)}}\right]\mathrm{Cov}\{\hat{\boldsymbol{\mu}}^{(k)}\}\left[\frac{\partial^2 D}{\partial \varphi \partial \boldsymbol{\mu}^{(k)}}\right]^T$$
$$= 2\cdot\Re\{\mathbb{T}\cdot\mathrm{Cov}\{\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\eta}}\}\cdot\mathbb{T}^H + \mathbb{T}\cdot\mathrm{Cov}\{\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\eta}}\}\cdot\mathbb{T}^T\} \tag{65}$$
$$= \cdot\Re\{2\mathbb{T}\}\cdot Cov\{\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\eta}}\}\cdot\Re\{2\mathbb{T}^T\},$$

where

$$\mathbb{T} = \sum_{l=1}^{M-N_{c1}}(\boldsymbol{v}_l^{(k)})^H\mathbf{T}^{(k)}\boldsymbol{\Upsilon}_l^{(k)}.$$

We rewrite $\boldsymbol{\Upsilon}_l^{(k)} = [\boldsymbol{\Upsilon}_{l,1}^{(k)}, \boldsymbol{\Upsilon}_{l,2}^{(k)}, \ldots, \boldsymbol{\Upsilon}_{l,N_{c1}N_{sum}}^{(k)}]$ with $\boldsymbol{\Upsilon}_{l,i}^{(k)} = [\boldsymbol{\Upsilon}_{l,i,\Re}^{(k)}, \boldsymbol{\Upsilon}_{l,i,\Im}^{(k)}]$. Correspondingly, we rewrite $\mathbb{T} = [\mathbb{T}_1, \mathbb{T}_2, \ldots, \mathbb{T}_{N_{c1}N_{sum}}]$ with $\mathbb{T}_i = [\mathbb{T}_{i,\Re}, \mathbb{T}_{i,\Im}]$. Then, we have $\mathbb{T}_{i,\Re} = \sum_{l=1}^{M-N_{c1}}(\boldsymbol{v}_l^{(k)})^H\mathbf{T}^{(k)}\boldsymbol{\Upsilon}_{l,i,\Re}^{(k)}$ and $\mathbb{T}_{i,\Im} = \sum_{l=1}^{M-N_{c1}}(\boldsymbol{v}_l^{(k)})^H\mathbf{T}^{(k)}\boldsymbol{\Upsilon}_{l,i,\Im}^{(k)}$. We further obtain

$$\mathbb{T}_{i,\Re} = \sum_{l=1}^{M-N_{c1}}(\boldsymbol{v}_l^{(k)})^H\mathbf{T}^{(k)}\left(-\frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \boldsymbol{\xi}}\left[\frac{\partial^2 G}{\partial \xi^2}\right]^{-1}\frac{\partial^2 G}{\partial \boldsymbol{\xi}\partial \boldsymbol{\eta}_i^\Re} + \frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \boldsymbol{\eta}_i^\Re}\right). \tag{66}$$

Note that

$$\sum_{l=1}^{M-N_{c1}}(\boldsymbol{v}_l^{(k)})^H\mathbf{T}^{(k)}\left(-\frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \boldsymbol{\xi}}\left[\frac{\partial^2 G}{\partial \xi^2}\right]^{-1}\frac{\partial^2 G}{\partial \xi \partial \boldsymbol{\eta}_i^\Re}\right)$$
$$= -\left[\frac{\partial^2 G}{\partial \xi^2}\right]^{-1}\sum_{l=1}^{M-N_{c1}}\sum_{z\neq l}^M \frac{(\boldsymbol{v}_l^{(k)})^H\mathbf{T}^{(k)}\boldsymbol{v}_z^{(k)}(\boldsymbol{v}_z^{(k)})^H\frac{\partial \boldsymbol{\Xi}^{(k)}}{\partial \xi}}{\lambda_l^{(k)} - \lambda_z^{(k)}}$$
$$2\Re\{\boldsymbol{e}_i^H\mathbf{Q}^{(k)}\}$$
$$= -\left[\frac{\partial^2 G}{\partial \xi^2}\right]^{-1}\sum_{l=1}^{M-N_{c1}}\sum_{z\neq l}^M \frac{(\boldsymbol{v}_l^{(k)})^H\mathbf{T}^{(k)}\boldsymbol{v}_z^{(k)}(\boldsymbol{v}_z^{(k)})^H\frac{\partial \boldsymbol{\Xi}^{(k)}}{\partial \xi}\boldsymbol{v}_l^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}} \tag{67}$$
$$\left(\boldsymbol{e}_i^H\mathbf{Q}^{(k)} + \boldsymbol{e}_i^T(\mathbf{Q}^{(k)})^T\right)$$
$$= -\left[\frac{\partial^2 G}{\partial \xi^2}\right]^{-1}\sum_{l=1}^{M-N_{c1}}\sum_{z=M-N_{c1}+1}^M \frac{(\boldsymbol{v}_l^{(k)})^H\mathbf{T}^{(k)}\boldsymbol{v}_z^{(k)}(\boldsymbol{v}_z^{(k)})^H\frac{\partial \boldsymbol{\Xi}^{(k)}}{\partial \xi}\boldsymbol{v}_l^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}}$$
$$\left(\boldsymbol{e}_i^H\mathbf{Q}^{(k)} + \boldsymbol{e}_i^T(\mathbf{Q}^{(k)})^T\right),$$

and

$$\sum_{l=1}^{M-N_{cl}} \left(\boldsymbol{v}_l^{(k)}\right)^H T^{(k)} \frac{\partial \boldsymbol{v}_l^{(k)}}{\partial \boldsymbol{\eta}_i^{\Re}}$$

$$= \sum_{l=1}^{M-N_{cl}} \sum_{z\neq l}^{M} \frac{\left(\boldsymbol{v}_l^{(k)}\right)^H T^{(k)} \boldsymbol{v}_z^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}} \sum_{p=1}^{N_k} e_i^H \boldsymbol{\Gamma}_p^{(k)} \boldsymbol{v}_l^{(k)} (\boldsymbol{v}_z^{(k)})^H (\boldsymbol{\Gamma}_p^{(k)})^H$$

$$+ (\boldsymbol{v}_z^{(k)})^H \boldsymbol{\Gamma}_p^H e_i (\boldsymbol{v}_l^{(k)})^T (\boldsymbol{\Gamma}_p^{(k)})^T \qquad (68)$$

$$= \sum_{l=1}^{M-N_{cl}} \sum_{z=M-N_{cl}+1}^{M} \frac{\left(\boldsymbol{v}_l^{(k)}\right)^H T^{(k)} \boldsymbol{v}_z^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}} \left(e_i^H \tilde{\boldsymbol{Q}}_{l,z}^{(k)} + e_i^T (\tilde{\boldsymbol{Q}}_{l,z}^{(k)})^T\right),$$

where $\tilde{\mathbf{Q}}_{l,z}^{(k)} = \sum_{p=1}^{N_k} \boldsymbol{\Gamma}_p^{(k)} \boldsymbol{v}_l^{(k)} (\boldsymbol{v}_l^{(k)})^H (\boldsymbol{\Gamma}_p^{(k)})^H$. Then, with (67) and (68), we can rewrite $\mathbb{T}_{i,\Re}$ as

$$\mathbb{T}_{i,\Re} = \boldsymbol{e}_i^H \mathbb{Q}^{(k)} + \boldsymbol{e}_i^T (\mathbb{Q}^{(k)})^T, \qquad (69)$$

where

$$\mathbb{Q}^{(k)} = \sum_{l=1}^{M-N_{cl}} \sum_{z=M-N_{cl}+1}^{M} \frac{(\boldsymbol{v}_l^{(k)})^H \mathbf{T}^{(k)} \boldsymbol{v}_z^{(k)}}{\lambda_l^{(k)} - \lambda_z^{(k)}} \left(\tilde{\mathbf{Q}}_{l,z}^{(k)} - \left[\frac{\partial^2 G}{\partial \xi^2}\right]^{-1} (\boldsymbol{v}_z^{(k)})^H \frac{\partial \Xi^{(\kappa)}}{\partial \boldsymbol{\xi}} \boldsymbol{v}_l^{(k)} \mathbf{Q}^{(k)}\right).$$

Likewise, we can also rewrite $\mathbb{T}_{i,\Im}$ as

$$\mathbb{T}_{i,\Im} = \mathbf{j} \cdot \left(\boldsymbol{e}_i^H \mathbb{Q}^{(k)} - \boldsymbol{e}_i^T (\mathbb{Q}^{(k)})^T\right). \qquad (70)$$

By defining $\tilde{\mathbb{Q}}^{(k)} = \mathbb{Q}^{(k)} + (\mathbb{Q}^{(k)})^H$, we obtain

$$\Re\{2\mathbb{T}_{i,\Re}\} = 2\Re\{\boldsymbol{e}_i^H \tilde{\mathbb{Q}}^{(k)}\}, \quad \Re\{2\mathbb{T}_{i,\Im}\} = -2\Im\{\boldsymbol{e}_i^H \tilde{\mathbb{Q}}^{(k)}\}. \quad (71)$$

Afterwards, following similar steps from (46) to (56), we can rewrite (65) as

$$\left[\frac{\partial^2 D}{\partial \varphi \partial \boldsymbol{\mu}^{(k)}}\right] \text{Cov}\{\hat{\boldsymbol{\mu}}^{(k)}\} \left[\frac{\partial^2 D}{\partial \varphi \partial \boldsymbol{\mu}^{(k)}}\right]^T$$

$$\simeq \frac{2}{L_s(N-L+1)^2} \Re \left[\sum_{t=0}^{N-L} \sum_{r=0}^{N-L} d_\sigma^T \left(\tilde{\boldsymbol{\Omega}}_{UV}^{(k)} \boldsymbol{\Delta}_{t,r,VV}^{(k)} \tilde{\boldsymbol{\Omega}}_{VU}^{(k)}\right) \right.$$

$$\left. \circ \left(\boldsymbol{\Delta}_{t,r,UU}^{(k)}\right)^*\right) d_\sigma\right] \qquad (72)$$

$$\simeq \frac{2\sigma_n^2}{\sigma_s^2 L_s(N-L+1)^2} \Re \left[\sum_{t=0}^{N-L} \sum_{r=0}^{N-L} \tilde{d}_\sigma^T \left(\left(\tilde{\boldsymbol{\Omega}}_{UV}^{(k)} \mathbf{V}_\gamma^H \mathbf{J}_{t,r} \mathbf{V}_\gamma \tilde{\boldsymbol{\Omega}}_{VU}^{(k)}\right) \right.\right.$$

$$\left.\left. \circ \left(\mathbf{U}_\gamma^H \mathbf{A} \boldsymbol{\Phi}^{t-r} \mathbf{A}^H \mathbf{U}_\gamma\right)^*\right) \tilde{d}_\sigma\right],$$

where

$$\tilde{\boldsymbol{\Omega}}_{UV}^{(k)} = \mathbf{U}_\gamma^H \tilde{\mathbb{Q}}^{(k)} \mathbf{V}_\gamma, \quad \tilde{\boldsymbol{\Omega}}_{VU}^{(k)} = \mathbf{V}_\Gamma^H \tilde{\mathbb{Q}}^{(k)} \mathbf{U}_\gamma.$$

Bearing in mind that

$$\frac{\partial^2 D^{(k)}}{\partial \varphi^2}\bigg|_{\varphi=\varphi_i^{(k)}} = 2(\boldsymbol{\alpha}(\varphi))^H \boldsymbol{\Psi}_1^H \left(\boldsymbol{v}^{(k)} (\boldsymbol{v}^{(k)})^H\right) \boldsymbol{\Psi}_1 \boldsymbol{\alpha}(\varphi)\bigg|_{\varphi=\varphi_i^{(k)}}, \quad (73)$$

we can then rewrite (59) as

$$\text{Cov}\{\hat{\varphi}_i^{(k)}\}$$

$$\simeq \frac{\sigma_n^2}{2\sigma_s^2 L_s(N-L+1)^2}$$

$$\cdot \frac{\Re\left[\sum_{t=0}^{N-L} \sum_{r=0}^{N-L} d_\sigma^{-T} \left(\left(\boldsymbol{\Omega}_{UV}(\mathbf{V}_\gamma^H \mathbf{J}_{t,r} \mathbf{V}_\gamma \boldsymbol{\Omega}_{VU}^{(k)}\right)^\circ \left(\mathbf{U}_\gamma^{HA} \boldsymbol{\Phi}^{t-r} \mathbf{A}^H \mathbf{U}_\gamma\right)^*\right) \tilde{d}_\sigma\right]}{\left[(\boldsymbol{\alpha}(\varphi))^H \boldsymbol{\Psi}_1^H \boldsymbol{v}^{(k)} (\boldsymbol{v}^{(k)})^H \boldsymbol{\Psi}_1 \boldsymbol{\alpha}(\varphi)\right]^2}\bigg|_{\varphi=\varphi_i^{(k)}}. \quad (74)$$

Interestingly, the DOA estimation variance in (74) exhibits a similar format with the CFO estimation variance given in (58). We can immediately observe that, the DOA estimation performance can be improved by increasing SNR or the number of blocks. Likewise, the relationship between the value of L and the DOA estimation performance is not that evident. We will leave its investigation in the simulations.

Computational complexity analysis

Let us now evaluate the computational complexity of our estimator in terms of the number of complex multiplications. We denote α as the number of trial CFO values to derive the solutions in (20). Calculation of $\hat{\mathbf{R}}_\gamma$ and its SVD requires $\mathcal{O}(M^2 N^2 L_s + M^3 L^3)$. For each trial CFO, calculation of $\boldsymbol{\Pi}^{(k)}(\xi)$ and its SVD requires $\mathcal{O}(M^2 N^2/2 + M^3)$. The number of total trial CFOs is αK. Then the complexity for the CFO estimation is in the order of $\mathcal{O}(M^2 N^2 L_s + M^3 L^3 + \alpha K(M^2 N^2/2 + M^3))$. On the side of DOA estimation for each user, a polynomial rooting with the highest order of $2M - 2$ is required whose complexity is in the order of $\mathcal{O}((2M-2)^3)$. Then, the overall required complexity for both CFO and DOA estimation for the K users is in the order of $\mathcal{O}(M^2 N^2 L_s + M^3 L^3 + \alpha K(M^2 N^2/2 + M^3) + K(2M-2)^3)$ For one case in our simulations, i.e., $N = 64$, $M = 4$, $L_s = 64$, $K = 8$, $L = 50$, and $\alpha = 140$,[a] the total complexity of our method is in the order of $\mathcal{O}(3.4 \times 10^7)$. The corresponding complexity of ESPRIT-2 is in the order of $\mathcal{O}(M^2 N^2 L_s + 3(M-1)^3 N^3) = (2.5 \times 10^7)$ [18]. Thus, the complexities of our method and ESPRIT-2 are in the same order. Bearing in mind that the algorithm is implemented at BS rather than the users, the required complexity is not the bottleneck [18].

Simulations

In this section, we assess the performance of the proposed CFO and DOA estimation algorithm from computer simulations. The total number of subcarriers is taken as $N = 64$. The quadrature phase-shift keying (QPSK) constellation is adopted. The number of multipaths within each channel cluster is $L_p = 10$. In each trial, the normalized CFO of each user is randomly

generated from -0.4 to 0.4, and all subcarriers are allocated to each user uniformly with randomly generated SAS. The root mean square error is adopted to evaluated the estimation performance.

First, we consider the single cluster scenario, i.e., $N_{cl} = 1$. Both the CFO and DOA estimation performance versus SNR are shown in Figure 1. In this example, we set $M = 2$, $K = 4$, $L_s = 64$ and $L = 50$, and the DOAs from set {30°, 45°, 95°, 100°}. Both the simulation and analytical results of the proposed method are presented in the figure. It is demonstrated that the analytical results closely match the simulations, especially under moderate or high SNR region, which verifies the correctness of our analysis. Besides, the performance of [18,19] are also presented for comparison, which are referred to as 'ESPRIT-1' and 'ESPRIT-2', respectively. Note that there is a similar smoothing parameter L in ESPRIT-1, which will be set the same as ours in the following unless otherwise stated. The results clearly demonstrate that substantial improvement can be achieved by our proposed method compared to the other two counterparts, especially in the lower SNR region. This can be explained as follows. First, we have exploited the smoothing technique in our method, which improves the estimation of the correlation matrix; Second

and more importantly, both ESPRIT-1 and ESPRIT-2 derive the CFO or DOA for each user by an indirect approach: they first estimate multiple CFO or DOA values separately for each occupied subcarrier of each user, and then combine them as the final result. Hence, lots of redundant parameters need to be estimated from the subspace algorithm, making both ESPRIT-1 and ESPRIT-2 less efficient. However, in our method, we make use of the fact that there is only one CFO and only one DOA for each user, and these two parameters are directly estimated by the designed rank reduction approach.

Next, we evaluate the performance of our method in the multi-cluster scenarios with $K = 4$ users. Both two clusters, i.e., $N_{cl} = 2$, and three clusters, i.e., $N_{cl} = 3$, are considered in this example. For $N_{cl} = 2$, the DOAs of the users are set as {{30°, 50°},{60°, 70°}, {90°, 140°}, {120°, 160°}}, while for $N_{cl} = 3$, the DOAs are set as {{30°, 50°, 140°}, {20°, 60°, 100°}, {90°, 110°, 140°}, {80°, 120°, 150°}. In addition, $M = 4$ and $M = 6$ are assumed for $N_{cl} = 2$ and $N_{cl} = 3$, respectively. Other parameters are set as $L_s = 64$ and $L = 50$. We depict the CFO and DOA estimation performance of our method in Figure 2, where both simulation and analytical curves are included. As expected, the results demonstrate the
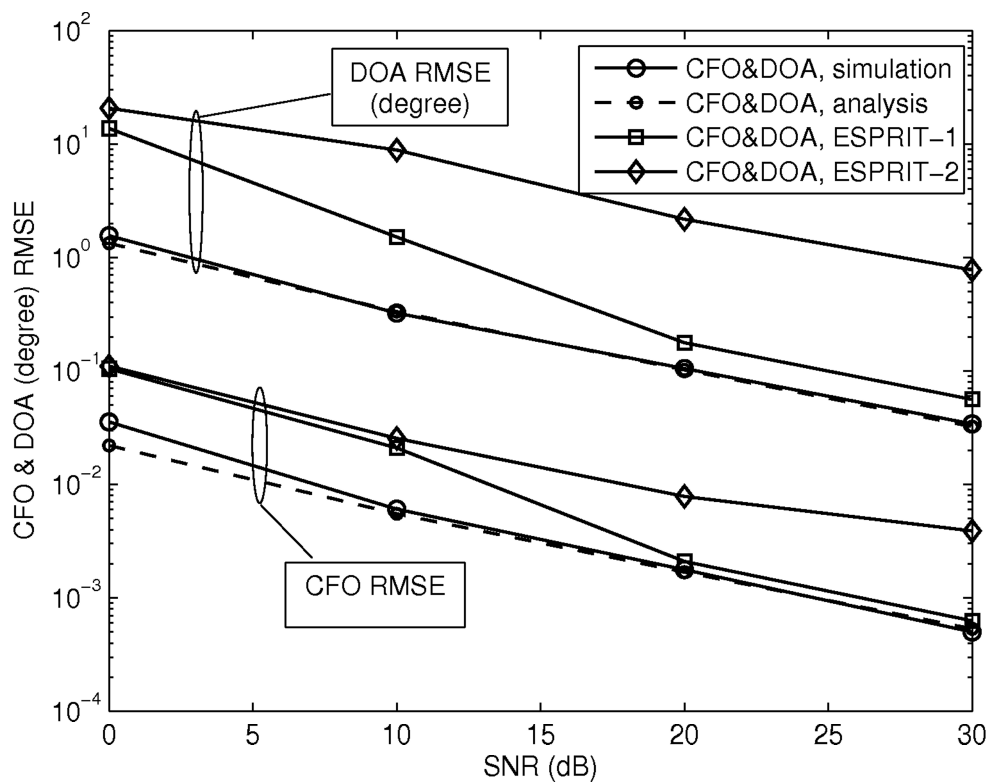

Figure 1 The performance comparison of both CFO and DOA estimation RMSE versus SNR under $N_{cl} = 1$ and $M = 2$. The top four curves correspond to the DOA RMSE (degree), while the bottom four curves correspond to the CFO RMSE. ○: simulation results of the proposed method; dashed: analytical results of the proposed method; □: the method from [19]; ◊: the method from [18].

effectiveness of our method for multi-cluster scenarios. The analytical results also closely match the simulations. Combing the observations from Figures 1 and 2, we can conclude that, as compared to ESPRIT-1 [19] and ESPRIT-2 [18], our method not only has the advantage of being applicable to multi-cluster channels, but also can obtain much better performance in single cluster channels.

In this example, we assume SNR = 20 dB, $M = 2$, $K = 4$, and $L_s = 64$, while the DOAs of the users are set the same as that in Figure 1. The estimation performance evolution of our method under the single cluster scenario is shown in Figure 3, with L increasing from the lower bound 49 calculated from (25)[b] to $N = 64$. The analytical results are presented by the dashed curves. For comparison, we also include the corresponding performance of ESPRIT-1 and ESPRIT-2. Note that since ESPRIT-2 does not have the parameter L, its performance is presented as a straight line parallel to the x-axis. From the results, we observe an optional range of L that can be selected without considerable loss of performance, e.g., from 49 to 62 in this example. Moreover, we see that, our method always obtains better DOA estimation performance than the other two candidates under all values of L. On the side of CFO estimation, our method always behaves better than ESPRIT-1, while as compared to ESPRIT-2, our method behaves worse only when $L = 64$. In addition, some mismatch appears between the simulation and analytical results of our method at the right-hand end of the curves. This is because that, when L is very large, the estimation for the correlation matrix $\mathbf{R}_\gamma$ becomes worse, whose fluctuation may result in the occurrence of some outliers in simulations.

Afterwards, we investigate the performance of our method with the increasing of L_s in Figure 4, where the DOA configuration is the same as that in Figure 3. Besides, we assume SNR = 20 dB, $M = 2$, $K = 4$, and $L = 50$ in this example. As expected, performance improvement can be observed with the increase of blocks adopted. Our method almost converges within 30 blocks. Moreover, it is seen that, in terms of both CFO and DOA estimation, our method behaves better than ESPRIT-1 under all values of L_s. Note that since ESPRIT-2 does not work when $L_s < 64$, its performance is not included in this figure. In addition, some mismatch appears between the analytical and simulation results of our method at the left-hand end of the curves. This is also because of the occurrence of outliers when a small number of blocks are adopted.
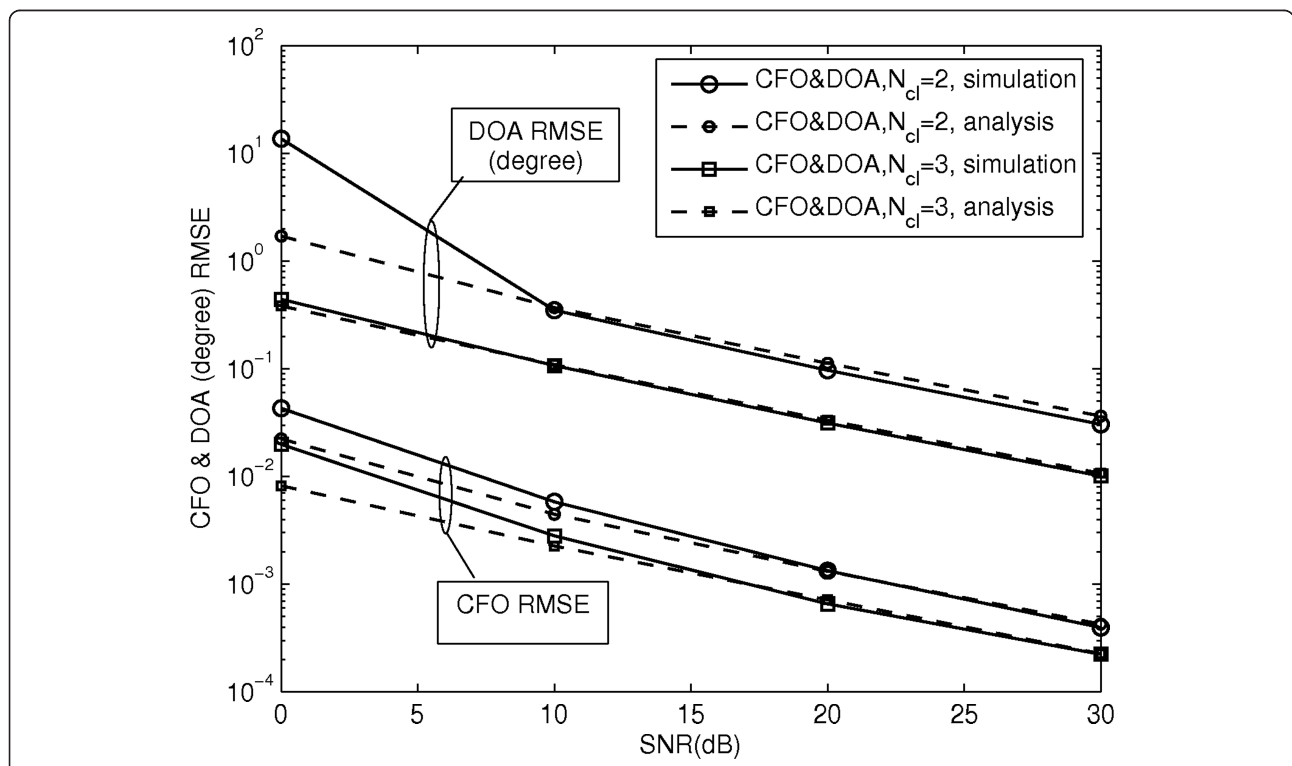

Figure 2 The CFO and DOA estimation RMSE performance versus SNR under multi-cluster scenarios. The top four curves correspond to the DOA RMSE (degree), while the bottom four curves correspond to the CFO RMSE. The solid and dashed curves denote the simulation and analytical results, respectively. ○: $N_{cl} = 2$ and $M = 4$; □: $N_{cl} = 3$ and $M = 6$.
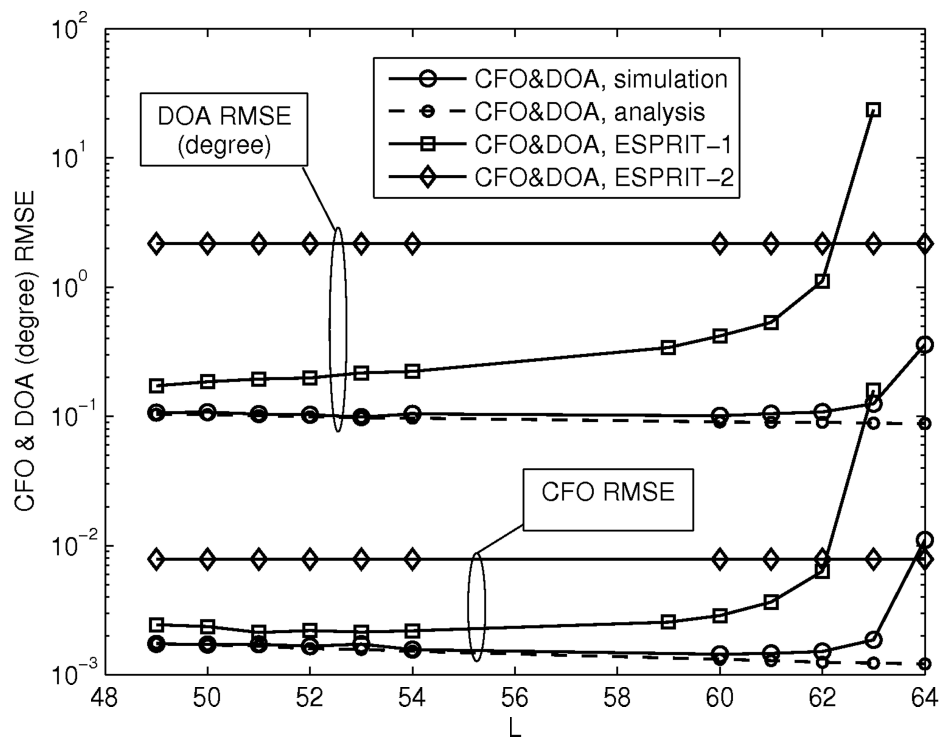
Figure 3 The CFO and DOA estimation RMSE performance versus *L* under the single cluster scenario. The top four curves correspond to the DOA RMSE (degree), while the bottom four curves correspond to the CFO RMSE. ○: simulation results of the proposed method; dashed: analytical results of the proposed method; □: the method from [19]; ◊: the method from [18].
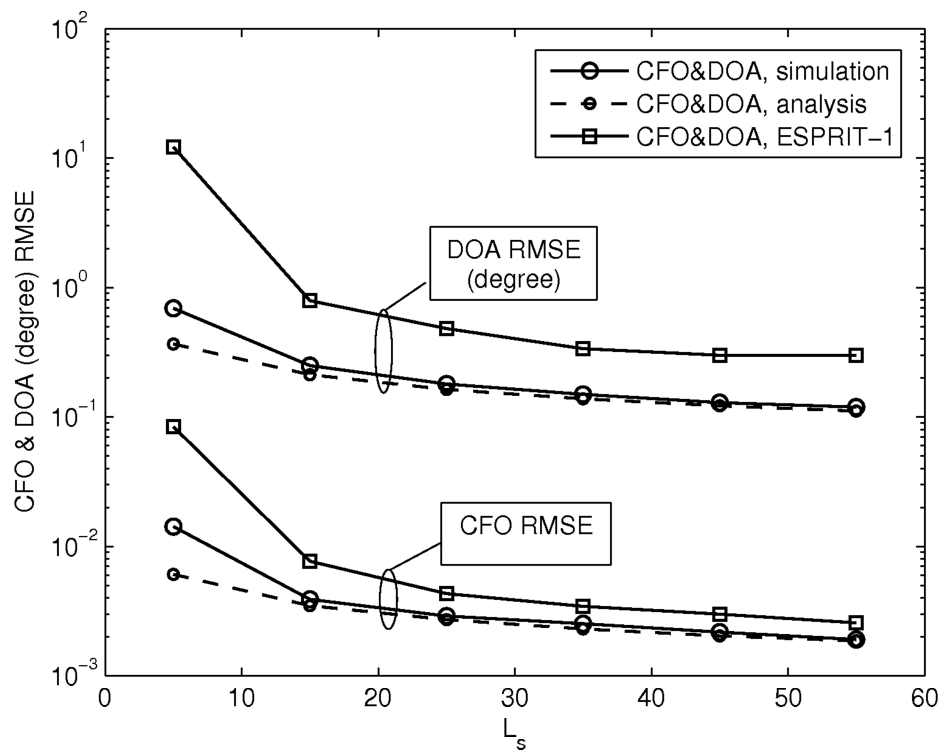

Figure 4 The CFO and DOA estimation RMSE performance versus *L_s* under the single cluster scenario. The top three curves correspond to the DOA RMSE (degree), while the bottom three curves correspond to the CFO RMSE.○: simulation results of the proposed method; dashed: analytical results of the proposed method; □: the method from [19].

In Figure 5, we consider the two-cluster scenario, i.e., $N_{\mathrm{cl}} = 2$, and investigate the estimation performance of our method under different number of users. In this example, we set $M = 4$, $L = 50$, $L_s = 64$. The DOAs of users are randomly picked out from set {{30°, 50°},{60°, 70°}, {90°, 140°}, {120°, 160°}, {140°, 170°}, {20°, 80°}, {50°, 100°}, {110°, 150°}}. For brevity, we only show the analytical results of $K = 8$ users in the figure. It is seen that, there is not much performance difference of our method with different number of users.

Conclusions

In this article, we developed a multiuser joint CFO and DOA estimation scheme for OFDMA uplink transmissions. We employed multi-antenna at the receiver and exploited the rank reduction approach to derive the CFO and multiple DOAs simultaneously for each user. The proposed scheme supports generalized SAS and can provide fully loaded transmissions such that the bandwidth efficiency is improved. Both analysis and simulation results corroborated the proposed study and demonstrate its advantages over the existing schemes.

Appendix 1

Proof of Lemma 1

First, it can readily be verified when $\boldsymbol{\omega} \in \mathrm{Span}(\boldsymbol{a}^{(k)})$, there holds $(\mathbb{B}_p^{(k)}(\xi^{(k)}) \otimes \boldsymbol{\omega})^H \mathbf{V}_{\boldsymbol{\gamma}} = 0$ Next for the case of $\boldsymbol{\omega} \notin \mathrm{Span}(\boldsymbol{a}^{(k)})$, a method of proof by contradiction is adopted. Note that $(\mathbb{B}_p^{(k)}(\xi^{(k)}) \otimes \boldsymbol{\omega})^H \mathbf{V}_{\boldsymbol{\gamma}} = 0$ indicates that $\mathbb{B}_p^{(k)}(\xi^{(k)}) \otimes \boldsymbol{\omega}$, i.e., $b_p^{(k)} \otimes \boldsymbol{\omega}$, belongs to the column space of $\mathbf{A}$. The implication is that there must exist $\boldsymbol{\varepsilon}_{p',i}^{(k')}$, not all zero, such that

$$
\begin{aligned}
\mathbf{b}_p^{(k)} \otimes \boldsymbol{\omega} &= \sum_{k'=1}^{K} \sum_{p'=1}^{N_{k'}} \sum_{i=1}^{N_{\mathrm{cl}}} \varepsilon_{p',i}^{(k')} (\mathbf{b}_{p'}^{(k')} \otimes \boldsymbol{a}_i^{(k')}) \\
&= \sum_{i=1}^{N_{\mathrm{cl}}} \varepsilon_p^{(k)} (\mathbf{b}_p^{(k)} \otimes \boldsymbol{a}_i^{(k)}) + \sum_{p' \neq p} \sum_{i=1}^{N_{\mathrm{cl}}} \varepsilon_{p',i}^{(k)} (\mathbf{b}_{p'}^{(k)} \otimes \boldsymbol{a}_i^{(k)}) \\
&\quad + \sum_{k' \neq k} \sum_{p'=1}^{N_{k'}} \sum_{i=1}^{N_{\mathrm{cl}}} \varepsilon_{p',i}^{(k')} (\mathbf{b}_{p'}^{(k')} \otimes \boldsymbol{a}_i^{(k')}).
\end{aligned}
\tag{75}
$$

An immediate observation is that when $\boldsymbol{\omega} \notin \mathrm{Span}(\boldsymbol{a}^{(k)})$, the coefficients $\boldsymbol{\varepsilon}_{p',i}^{(k)}$ and $\boldsymbol{\varepsilon}_{p',i}^{(k')}$ of the last two terms in (75) cannot all equal zeros. On the other hand, we can rewrite (75) into the following two equivalent equations:

$$
\underline{\mathbf{b}}_p^{(k)} \otimes \boldsymbol{\omega} = \sum_{k'=1}^{K} \sum_{p'=1}^{N_{k'}} \sum_{i=1}^{N_{\mathrm{c}1}} {}_{p,i}^{(k')} (\underline{\mathbf{b}}_{p'}^{(k')} \otimes \mathbf{a}_i^{(k')})
\tag{76}
$$

$$
\mathrm{e}^{\mathrm{j}2\pi\theta_p^{(k)}} (\underline{\mathbf{b}}_p^{(k)} \otimes \boldsymbol{\omega}) = \sum_{k'=1}^{K} \sum_{p'=1}^{N_{k'}} \sum_{i=1}^{N_{\mathrm{c}1}} \varepsilon_{p',i}^{(k')} \mathrm{e}^{\mathrm{j}2\pi\theta_{p'}^{(k')}} (\underline{\mathbf{b}}_{p'}^{(k')} \otimes \mathbf{a}_i^{(k')}),
\tag{77}
$$

Figure 5 The CFO and DOA estimation RMSE performance versus SNR with different number of users. The top four curves correspond to the DOA RMSE (degree), while the bottom four curves correspond to the CFO RMSE. ○: $K = 2$; □: $K = 4$; ◊: $K = 8$.
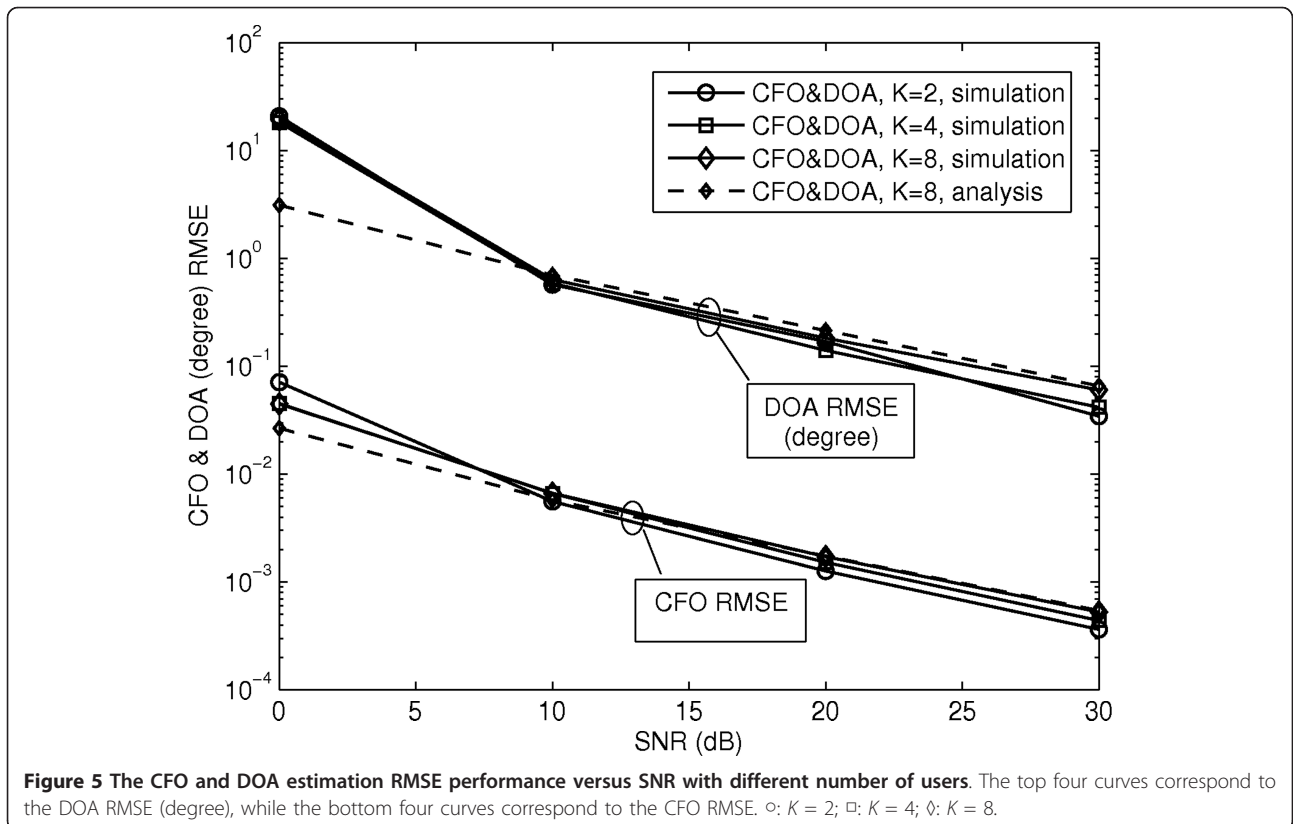
where $\underline{\mathbf{b}}_p^{(k)}$ has been defined in (19). Combing the above two equations, i.e., $e^{j2\pi\Theta_p^{(k)}} \times (76) - (77)$, yields

$$\sum_{p'\neq p}\sum_{i=1}^{N_{cl}}\varepsilon_{p',i}^{(k)}(e^{j2\pi\Theta_p^{(k)}} - e^{j2\pi\Theta_{p'}^{(k)}})(\underline{\mathbf{b}}_{p'}^{(k)} \otimes \mathbf{a}_i^{(k)})$$
$$+ \sum_{k'\neq k}\sum_{p'=1}^{N_{k'}}\sum_{i=1}^{N_{cl}}\varepsilon_{p',i}^{(k)}(e^{j2\pi\Theta_p^{(k)}} - e^{j2\pi\Theta_{p'}^{(k')}})(\underline{\mathbf{b}}_{p'}^{(k')} \otimes \mathbf{a}_i^{(k')}) = 0. \tag{78}$$

Note that above $(N_k - 1)N_{cl}$ vectors $\underline{\mathbf{b}}_{p'}^{(k)} \otimes \mathbf{a}_i^{(k)}$ and $\sum_{k'=1}^{K}N_{c1}N_{k'}$ vectors $\underline{\mathbf{b}}_{p'}^{(k')} \otimes \mathbf{a}_i^{(k')}$ are column vectors of $\underline{\mathbf{A}}$. When $\underline{\mathbf{A}}$ has full column rank, these vectors will be uncorrelated with each other. Then the coefficients $\boldsymbol{\varepsilon}_{p',i}^{(k)}$ and $\boldsymbol{\varepsilon}_{p',i}^{(k')}$ in (78) should all equal zeros, which is in contrary with the observation from (75). Hence, $\underline{\mathbf{b}}_{p'}^{(k)} \otimes \omega$ does not belong to the column space of $\underline{\mathbf{A}}$, which implies $(\mathbb{B}_p^{(k)}(\xi^{(k)}) \otimes \boldsymbol{\omega})^H \mathbf{V}_\gamma \neq \mathbf{0}$ when $\omega \notin \text{Span}(\boldsymbol{a}^{(k)})$.

Likewise, we can prove (18) as follows. Assume $\mathbb{B}_p^{(k)}(\xi) \otimes \boldsymbol{\omega}$ belongs to the column space of $\mathbf{A}$, then there exists $\boldsymbol{\varepsilon}_{p',i}^{(k')}$, not all zero, such that

$$\mathbb{B}_p^{(k)}(\xi) \otimes \boldsymbol{\omega} = \sum_{k'=1}^{K}\sum_{p'=1}^{N_{k'}}\sum_{i=1}^{N_{c1}}\varepsilon_{p',i}^{(k')}(\mathbf{b}_{p'}^{(k')} \otimes \mathbf{a}_i^{(k')}). \tag{79}$$

Further, we rewrite the above equation as

$$\mathbb{B}_p^{(k)}(\xi) \otimes \boldsymbol{\omega} = \sum_{k'=1}^{K}\sum_{p'=1}^{N_{k'}}\sum_{i=1}^{N_{c1}}\varepsilon_{p',i}^{(k')}(\mathbf{b}_{p'}^{(k')} \otimes \mathbf{a}_i^{(k')}), \tag{80}$$

$$e^{j\frac{2\pi(c_p^{(k)}+\xi)}{N}}(\mathbb{B}_p^{(k)}(\xi) \otimes \boldsymbol{\omega})$$
$$= \sum_{k'=1}^{K}\sum_{p'=1}^{N_{k'}}\sum_{i=1}^{N_{cl}}\varepsilon_{p',i}^{(k')}e^{j2\pi\theta_{p'}^{(k')}}(\underline{\mathbf{b}}_{p'}^{(k')} \otimes \mathbf{a}_i^{(k')}). \tag{81}$$

Combing these two equations yields

$$\sum_{k'=1}^{K}\sum_{p'=1}^{N_{k'}}\sum_{i=1}^{N_{c1}}\varepsilon_{p',i}^{(k')}(e^{j2\pi\Theta_{p'}^{(k')}} - e^{j\frac{2\pi(c_p^{(k)}+\xi)}{N}})(\underline{\mathbf{b}}_{p'}^{(k')} \otimes \mathbf{a}_i^{(k')}) = 0. \tag{82}$$

When $\xi \neq \xi^{(k)}$, we have $e^{j2\pi\Theta_{p'}^{(k')}} \neq e^{j\frac{2\pi(c_p^{(k)}+\xi)}{N}}$ for all k' and p'. As a result, when $\underline{\mathbf{A}}$ has full column rank, all coefficients $\boldsymbol{\varepsilon}_{p',i}^{(k')}$ should equal to zeros. Thus, this is in

contrary with the assumption of (79). The implication is that $\mathbb{B}_p^{(k)}(\xi) \otimes \omega$ does not belong to the column space of $\mathbf{A}$ when $\xi \neq \xi^{(k)}$. Finally, we arrive at (18).

Appendix 2
Proof of Lemma 2
Define $\nu = \lfloor M/N_{cl} \rfloor$ and assume $\nu \geq 2$. Here, we consider only the case of $M = \nu N_{cl}$. Note that the results obtained here is directly applicable to the case of $M > \nu N_{cl}$.

By defining $M \times M$ full rank matrix

$$\mathbf{a}^{(1,\nu)} = [\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(\nu)}], \tag{83}$$

we obtain

$$\mathbf{a}^{(k)} = \mathbf{a}^{(1,\nu)}\mathbf{c}^{(k)}, \quad k > \nu, \tag{84}$$

where the $M \times N_{cl}$ matrix $\mathbf{c}^{(k)}$ has full column rank. Further, there is

$$\underline{\mathbf{B}}^{(k)} \otimes \mathbf{a}^{(k)} = \underline{\mathbf{B}}^{(k)} \otimes (\mathbf{a}^{(1,\nu)}\mathbf{c}^{(k)}) = (\underline{\mathbf{B}}^{(k)} \otimes \mathbf{a}^{(1,\nu)})(I_{N_k} \otimes \mathbf{c}^{(k)}), \tag{85}$$

which implies that the column space of $\mathbf{B}^{(k)} \otimes \mathbf{a}^{(k)}$ belongs to the column space of $\mathbf{B}^{(k)} \otimes \mathbf{a}^{(1,\nu)}$ Then, we define a matrix

$$\underline{\tilde{\mathbf{A}}} = \left[\underline{\mathbf{B}}^{(1)} \otimes \boldsymbol{a}^{(1)}, \underline{\mathbf{B}}^{(2)} \otimes \boldsymbol{a}^{(2)}, \ldots, \underline{\mathbf{B}}^{(\nu)} \otimes \boldsymbol{a}^{(\nu)}, \right.$$
$$\left.\left[\underline{\mathbf{B}}^{(\nu+1)}, \underline{\mathbf{B}}^{(\nu+2)}, \ldots, \underline{\mathbf{B}}^{(K)}\right] \otimes \boldsymbol{a}^{(1,\nu)}\right]$$
$$= \left[\underline{\tilde{\mathbf{B}}}^{(1)} \otimes \boldsymbol{a}^{(1)}, \underline{\tilde{\mathbf{B}}}^{(2)} \otimes \boldsymbol{a}^{(2)}, \ldots, \underline{\tilde{\mathbf{B}}} \otimes \boldsymbol{a}^{(\nu)}\right] \tag{86}$$

where

$$\underline{\tilde{\mathbf{B}}}^{(k)} = [\underline{\mathbf{B}}^{(k)}, \underline{\mathbf{B}}^{(\nu+1)}, \ldots, \underline{\mathbf{B}}^{(K)}]. \tag{87}$$

An important observation is that when $\tilde{\mathbf{A}}$ has full column rank, $\underline{\mathbf{A}}$ has full column rank too. Afterwards, since the disjointness among $\boldsymbol{a}^{(k)}$, $k = 1, 2, \ldots, \nu$, we know

$$\text{rank}(\underline{\tilde{\mathbf{A}}}) = \sum_{k=1}^{\nu}\text{rank}(\underline{\tilde{\mathbf{B}}}^{(k)} \otimes \boldsymbol{a}^{(k)})$$
$$= \sum_{k=1}^{\nu}N_{cl} \cdot \min(L - 1, N_k + \sum_{\nu+1}^{K}N_i). \tag{88}$$

Clearly, when

$$L - 1 \geq \max_{1\leq k\leq\nu}(N_k + \sum_{\nu+1}^{K}N_i), \tag{89}$$

i.e.,

$$L \geq 1 + \max_{1\leq k\leq\nu}(N_k) + \sum_{\nu+1}^{K}N_i, \tag{90}$$

$\underline{\tilde{\mathbf{A}}}$ has full column rank. Furthermore, considering the arbitrariness of user arrangement, L only needs to be not less than the minimum value of the right-hand side of (90). Finally, (25) can be directly derived from above observations.

Endnotes

[a] In our simulations, we first obtain a coarse CFO estimation value by searching over (-0.5, 0.5) with an interval of 10^{-2}, denoted by $\hat{\xi}_c^{(k)}$. Then, the finer estimation $\hat{\xi}_f^{(k)}$ is achieved by searching over $(\hat{\xi}_c^{(k)} - 0.01, \hat{\xi}_c^{(k)} + 0.01)$ with an interval of 10^{-3}. The final estimation $\hat{\xi}^{(k)}$ is obtained by searching over $(\hat{\xi}_f^{(k)} - 10^{-3}, \hat{\xi}_f^{(k)} + 10^{-3})$ using an interval of 10^{-4}. Thus, the number of total trial CFO values is $\alpha = 140$. [b] From (25), we know $\nu = \left\lfloor \frac{M}{N_{c1}} \right\rfloor = 2 < K = 4$ in this example. Then, according to (25), we have $L \geq N_1 + 1 + \sum_{k=3}^{4} N_k = 49$.

Abbreviations

1D: one-dimensional; CFO: carrier frequency offset; CP: cyclic prefix; DOA: direction-of-arrival; ESPRIT: estimation of signal parameters via rotational invariance; ML: maximum likelihood; QPSK: quadrature phase-shift keying; SAS: subcarrier assignment scheme; SVD: singular value decomposition.

Author details

[1] MOE Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, P. R. China [2] Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, 100084, P. R. China [3] National Mobile Communications Research Laboratory, Southeast University, 210096, P. R. China

Competing interests

The authors declare that they have no competing interests.

References

1. GL Stuber, JR Barry, SW Mclaughlin, Y Li, MA Ingram, TG Pratt, Broadband MIMO-OFDM wireless communications. Proc IEEE. **92**, 271–294 (2004). doi:10.1109/JPROC.2003.821912
2. M Morelli, Timing and frequency synchronization for the uplink of an OFDMA system. IEEE Trans Commun. **52**, 292–306 (2004)
3. M Morelli, C Kuo, M Pun, Synchronization techniques for orthogonal frequency division multiple access (OFDMA): a tutorial review. Proc IEEE. **95**, 1394–1427 (2007)
4. P Sun, L Zhang, Low complexity pilot aided frequency synchronization for OFDMA uplink transmission. IEEE Trans Wirel Commun. **8**, 3758–3769 (2009)
5. MR Raghavendra, E Lior, S Bhashyam, K Giridhar, Parametric channel estimation for pseudorandom tile-allocation in uplink OFDMA. IEEE Trans Signal Process. **55**, 5307–5381 (2007)
6. MO Pun, M Morelli, CCJ Kuo, Maximum-likelihood synchronization and channel estimation for OFDMA uplink transmissions. IEEE Trans Commun. **54**, 726–736 (2006)
7. Z Wang, Y Xin, G Mathew, Carrier-frequency offset estimation for OFDMA uplink with generalized subcarrier-assignment, in Proc of ICC, Beijing, China, 3490–3494 (2008)
8. Z Wang, Y Xin, G Mathew, Iterative carrier-frequency offset estimation for generalized OFDMA uplink transmission. IEEE Trans Wirel Commun. **8**, 1373–1383 (2009)
9. S Kay, S Saha, Mean likelihood frequency estimation. IEEE Trans Signal Process. **48**, 1937–1946 (2000). doi:10.1109/78.847780
10. J Chen, YC Wu, SC Chan, TS Ng, Joint maximum-likelihood CFO and channel estimation for OFDMA uplink using importance sampling. IEEE Trans Veh Technol. **57**, 3462–3470 (2008)
11. S Sezginer, P Bianchi, Asymptotically efficient reduced complexity frequency offset and channel estimators for uplink MIMO-OFDMA systems. IEEE Trans Signal Process. **56**, 964–979 (2008)
12. S Barbarossa, M Pompili, GB Giannakis, Channel-independent synchronization of orthogonal frequency division multiple access systems. IEEE J Sel Areas Commun. **20**, 474–486 (2002). doi:10.1109/49.983375
13. Z Cao, U Tureii, YD Yao, Deterministic multiuser carrier-frequency offset estimation for interleaved OFDMA uplink. IEEE Trans Commun. **52**, 1585–1594 (2004). doi:10.1109/TCOMM.2004.833183
14. RO Schmidt, Multiple emitter location and signal parameter estimation, in Proc RADC Spectral Estimation Workshop, Rome, NY, USA, 242–258 (1979)
15. J Lee, S Lee, KJ Bang, S Cha, D Hong, Carrier frequency offset estimation using ESPRIT for interleaved OFDMA uplink systems. IEEE Trans Veh Technol. **56**, 3227–3231 (2007)
16. Y Zhu, KB Letaief, CFO estimation and compensation in single carrier interleaved FDMA systems, in Proc IEEE GLOBECOM, Honolulu, HI, USA, 1–5 (2009)
17. KH Wu, WH Fang, YT Chen, Joint carrier frequency offset and direction of arrival estimation via hierarchical ESPRIT for interleaved OFDMA/SDMA uplink systems, in Proc IEEE VTC, Taipei, Taiwan, China, Spring, 1–5 (2010)
18. H Wang, Q Yin, Multiuser carrier frequency offsets estimation for OFDMA uplink with generalized carrier assignment scheme. IEEE Trans Wirel Commun. **8**, 3347–3353 (2009)
19. W Zhang, Q Yin, W Wang, Carrier frequency offset estimation for OFDMA uplink using multi-antenna, in Proc IEEE ICC, Cape Town, South Africa, 1–5 (2010)
20. H Asplund, AF Molisch, M Steinbauer, NB Mehta, Clustering of scatterers in mobile radio channels-evaluation and modeling in the COST259 directional channel model, in Proc IEEE ICC, New York City, USA, **2**, 901–905 (2002)
21. J Fuhl, A Molisch, E Bonek, Unified channel model for mobile radio systems with smart antennas. IEE Proc Radar, Sonar and Navigation. **145**, 32–41 (1998). doi:10.1049/ip-rsn:19981750
22. F Gao, A Nallanathan, Blind maximum likelihood CFO estimation for OFDM systems via polynomial rooting. IEEE Signal Process Lett. **13**, 73–76 (2006)
23. H Clergeot, S Tressens, A Ouamri, Performance of high resolution frequencies estimation methods compared to the CramerRao bounds. IEEE Trans ASSP. **11**, 1703–1720 (1989)
24. BD Rao, KVS Hari, Effect of spatial smoothing on the performance of MUSIC and the minimum-norm method. IEE Proc. **137**, 449–458 (1990)